

THIS WEEK

EDITORIALS

FOOD Fertilizer is the best way to feed Africa right now **p.510**



WORLD VIEW Teach young scientists how to manage their labs **p.511**

VENICE City of Water edges closer to becoming a city under water **p.512**

Must try harder

Too many sloppy mistakes are creeping into scientific papers. Lab heads must look more rigorously at the data — and at themselves.

Science: Branch of knowledge or study dealing with a body of facts or truths systematically arranged. So says the dictionary. But, as most scientists appreciate, the fruits of what is called science are occasionally anything but. Most of the time, when attention focuses on divergence from this gold (and linguistic) standard of science, it is fraud and fabrication — the facts and truth — that are in the spotlight. These remain important problems, but this week *Nature* highlights another, more endemic, failure — the increasing number of cases in which, although the facts and truth have been established, scientists fail to make sure that they are systematically arranged. Put simply, there are too many careless mistakes creeping into scientific papers — in our pages and elsewhere.

A Comment article on page 531 exposes one possible impact of such carelessness. Glenn Begley and Lee Ellis analyse the low number of cancer-research studies that have been converted into clinical success, and conclude that a major factor is the overall poor quality of published pre-clinical data. A warning sign, they say, should be the “shocking” number of research papers in the field for which the main findings could not be reproduced. To be clear, this is not fraud — and there can be legitimate technical reasons why basic research findings do not stand up in clinical work. But the overall impression the article leaves is of insufficient thoroughness in the way that too many researchers present their data.

The finding resonates with a growing sense of unease among specialist editors on this journal, and not just in the field of oncology. Across the life sciences, handling corrections that have arisen from avoidable errors in manuscripts has become an uncomfortable part of the publishing process.

The evidence is largely anecdotal. So here are the anecdotes: unrelated data panels; missing references; incorrect controls; undeclared cosmetic adjustments to figures; duplications; reserve figures and dummy text included; inaccurate and incomplete methods; and improper use of statistics — the failure to understand the difference between technical replicates and independent experiments, for example.

It is usually the case that original data can be produced, mistakes corrected, and the findings of the corrected research paper still stand. At the very least, however, there is too little attention paid and too many corrections, which reflect unacceptable shoddiness in laboratories that risks damaging trust in the science that they, and others, produce.

The situation throws up many questions. Here are three of them. Who is responsible? Why is it happening? How can it be stopped?

The principal investigators (PIs) of any lab from which the work originates, especially if their names are on the paper, have an absolute and unavoidable responsibility to ensure the quality of the data from their labs, even if the main work is done by experienced postdocs. Officially, postdocs and graduate students are still in training, and it is the PI's job to make sure they are properly trained — in statistics and appropriate image editing, for a start. It is unacceptable for lab heads — who are happy to take the credit for good work — to look at raw data

for the first time only when problems in published studies are reported.

In private, scientists who run labs in even the most prestigious universities admit that they have little time to supervise and train all their students. Institutions such as the European Molecular Biology Laboratory in Heidelberg, Germany, have maximum lab sizes for this reason. Funding agencies should require grant applicants to indicate lab size and offer adequate supervision. As is the case in commercial compa-

“Handling corrections that have arisen from avoidable errors in manuscripts has become an uncomfortable part of the publishing process.”

nies, larger labs should introduce formal training and a management hierarchy, with more experienced postdocs and research associates required to sign off data and experiments if PIs cannot do so themselves.

What can journal editors and referees do? Sloppiness is sometimes caught, but so much must be taken on trust. Journals should certainly offer online commenting, so that alert readers can point out errors. Where comments or corrections appear in other journals, these should be linked

from the original paper — as the Comment authors recommend.

There should also be increased scope to publish fuller results from an experiment, and subsequent negative or positive corroborations. There is an opportunity here for ‘minimum threshold’ journals, such as *PLoS ONE* and *Scientific Reports*. Editors and referees cannot be expected to divine when only positive data are included and inconvenient results left out, but journals should encourage online presentation of the complete picture. And scientists should offer it. The complete picture is, after all, what this science of ours strives to provide. ■

Under surveillance

Global systems for monitoring threats from flu need a radical overhaul.

Imagine a global weather and climate forecasting system that collects data regularly in just a handful of countries, and takes measurements elsewhere only during extreme weather events. That is what today's global flu-surveillance system mostly looks like.

The shortcomings of flu surveillance have long been recognized (see *Nature* 440, 6–7; 2006), but they are attracting renewed attention following the creation in labs of strains of the H5N1 avian influenza virus that can spread between mammals. The main cited public-health benefit of the research is that it will allow for monitoring for such mutations in the wild, and give a remote chance of containing an emerging pandemic.

It is certainly urgent to monitor wild flu strains for mutations that might make them transmissible between mammals (see *Nature* **482**, 439; 2012). But as Malik Peiris, a flu virologist at the University of Hong Kong, says, detection of a breaking pandemic is “a very ambitious goal, and this is where vastly enhanced global surveillance is needed”.

Current surveillance can barely identify threats, let alone track them. The precursor to the H1N1 virus that caused a pandemic in 2009 had been circulating worldwide for years in pigs, and the pandemic virus had been infecting humans in Mexico for months, before either was detected. That virus is also a reminder that threats come from many flu subtypes other than H5N1.

An analysis by *Nature* shows that timely, continued and representative global surveillance of the genetic sequences of flu isolates from pigs and poultry just isn't happening (see page 520). From 2003 to 2011, most countries collected few or no sequences, and genetic surveillance of flu in pigs was and is almost non-existent. There is typically a lag of years between collection of viruses and the release of their sequences into public databases, so there are very few data on their recent evolution.

Yet the analysis gives hope that this situation could be rectified, given political will, modest funding and international coordination. Hong Kong has collected the most flu sequences from pigs after the United States and China, and most of those come from labs at the University of Hong Kong, including Peiris's; this shows what a few dedicated centres can achieve. Similarly, the Influenza Genome Sequencing Project of the US National Institute of Allergy and Infectious Diseases, which was launched in 2004 and sequences whole flu genomes from isolates collected globally, accounts for around half of sequences generated

worldwide. And in the past decade, many nations affected by H5N1 have greatly improved their surveillance, often despite limited resources and poor veterinary and health infrastructure.

More sequencing alone is not enough. Sequences tend currently to come in fits and starts, in response to an outbreak, one-off projects or as funding allows, and there is little sustained passive surveillance. Global, scientific and representative sampling is needed, from multiple outbreaks and diverse populations, taking into account risk factors such as the size of livestock populations, husbandry practices and proximity to waterfowl reservoirs.

Funding is not the only problem. Few countries, for example, compensate for culled animals to encourage farmers to report outbreaks; and some might conceal, or not actively look for, flu infections for trade reasons. Nations can be reluctant to share viral isolates if they do not get anything in return, although the World Health Organization's Pandemic Influenza Preparedness Framework, published last year, should help to ensure that they do get appropriate benefits, including access to vaccines.

Surveillance makes sense even without the promise of tracking a pandemic. Detecting outbreaks in livestock allows control through culling or vaccination to avoid crippling losses, and limits the opportunities for viruses to mutate, outpace vaccines and possibly turn pandemic. Surveillance also generates crucial data for epidemiology and drug-resistance monitoring, yet it remains a low priority. Sequencing costs can fall all they like, but without greater, and more sustained, routine surveillance efforts, there will be few samples to sequence. ■

Food for thought

In the short term, chemical fertilizers are the best way to feed Africa.

Chemical fertilizers get a bad press, with some justification. Their use can pollute water supplies and generate significant greenhouse-gas emissions. But they are an excellent way to boost crop yields: they help to grow food. And in sub-Saharan Africa that means they can help to fill empty bellies and save lives.

Parts of Africa sorely need that help. Across the continent, agricultural lands are characterized by red soil that is low in nutrients. Intensive farming has seen the typical hectare of sub-Saharan farmland lose 22 kilograms of nitrogen, 2.5 kg of phosphorus and 15 kg of potassium annually over the past 30 years. Impoverished African farmers cannot afford to wait for the international community to deliberate on the long-term, green methods needed for a sustainable global agricultural system. They need to deploy methods that work now — and that means that in the short term, they need access to chemical fertilizers.

But there lies the problem. Most of the Malawian farmers interviewed for the Feature on page 525, for example, said that their biggest problem is the high cost and poor supply of fertilizer.

Although this evidence is anecdotal, it hints at something more. Farming a smallholding is intensive, backbreaking work that, for the most part, is done out of necessity, not choice. Greener practices such as no-till farming may be cheaper than using fertilizer, but they are less efficient and lack appeal because they add to farmers' already hard labour. By contrast, the quick and easy gains of fertilizers free up farmers' time, and can turn a subsistence existence into a commercial operation, offering a potential way to escape the crushing cycle of poverty.

Africa is not a laboratory in which to investigate and promote alternative agricultural strategies at the expense of those that are known to work. Development efforts need instead to focus first on the problems

of fertilizer supply and cost. Without fertilizers, it is hard to see how African farmers can catch up with their counterparts in Europe, North America and Asia, all of whom benefited from the boost chemicals gave their agricultural enterprise.

Subsidies are a key way to tackle this problem. It is encouraging that economists, including some at the World Bank, are recognizing that subsidies can help private-sector development, as demonstrated by Rwanda's agriculture-support programme, which meets some of the costs of fertilizer transport. Still, last week's *coup d'état* in Mali is a worrisome reminder of the political volatility in some parts of the continent, which can make donors reluctant to inject more cash because they fear that it will not reach its intended target.

There are other financial tools, too. Kenya's Equity Bank, for example, set up a loan system in 2008 to help farmers buy fertilizer. It lends the farmers cash at the start of the agricultural season, when they need fertilizer most but are least able to afford it, and allows them to pay back the loan at harvest time, when they sell their crop.

Improved access to fertilizer, although essential at present, is not the best long-term solution. Research must continue to reduce reliance on chemicals and to make their use more efficient. As highlighted on page 525, accurate and detailed information being gathered on soil types and health can allow for more precise and appropriate fertilizer application. Work is also needed to reduce the thirst of crops for fertilizer, and further improvements could be made by manipulating the soil and microorganisms around plant roots to increase the amount of dissolved nitrogen available to the crops.

Ultimately, the move towards a more sustainable form of agriculture will require investment to help farmers in Africa take advantage of alternatives to chemical fertilizers, although they will need to be convinced of the benefits of these new approaches.

The key to success is for farmers to choose the practices that will work best for them — past development efforts have shown that without buy-in from local communities, initiatives simply don't work. For now, that has to mean improved access to fertilizers, because the choice between food and famine is an easy one. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunjv



Scientists must be taught to manage

Young scientists need more help to set up and run research labs, says Jessica C. Seeliger.

Starting an academic lab is like launching a small business. But does scientific training really prepare us for success? As a young investigator just over a year into my job, I feel pressure — much of it self-generated — to produce results, attract funding and ultimately to make a name for myself in my chosen field of bacterial pathogenesis.

As researchers, we are trained to work within a rational and methodical framework. But when it comes to running our labs and managing people, we have to rely on our gut feelings, our limited know-how from mentoring a few students or our observations of our previous advisers. We can often feel ill-prepared.

Take dealing with a difficult co-worker or motivating students. As scientists, we must be honest with someone about faults in data or reasoning. But while striving for this scientific objectivity, we can forget the importance of body language and of directing discussion at a problem rather than a person. And even something as apparently straightforward as having a meeting can be problematic. The many collective hours spent around conference tables can feel like lost time when agendas wander and goals are not met.

Would we do any better if we received formal training that gave us a logical framework for lab management? Some young investigators would no doubt argue that such training is inefficient and ineffective. The classic method is to work from your own experience in your mentors' labs. Although this is a valuable starting point, building a new lab and serving as its sole head is a very different prospect from working in an established lab with senior students and support staff. So my current support network consists mainly of a handful of other young investigators, all of us amazed by the universality of the challenges we face. We trade tips and anecdotes about recruiting and retaining, motivating and negotiating, and we agonize over mistakes.

So, we need help — or at least, some of us do. Yet funding agencies offer no routine management training for people at my level. This is despite the many career-progression programmes and workshops now available for graduate students and postdocs.

The Burroughs Wellcome Fund and Howard Hughes Medical Institute did create a course for people at my stage of a scientific career, called 'Making the Right Moves'. But the course ran only twice — in 2002 and 2005. What endures is a book based on the course, which, along with Kathy Barker's *At the Helm* and *Lab Dynamics* by Carl Cohen and Suzanne Cohen, constitutes almost the entire reference library available to new investigators.

Recognition of this training void has come recently from an unexpected corner: the American Express Foundation, which last year started

to fund an annual 'Workshop on Leadership in BioScience' at Cold Spring Harbor Laboratory in New York.

Last month I went on the course, alongside my husband — Markus Seeliger, also a young investigator — and 25 scientists from around the world at a similar stage of their careers, for three days of lectures, role-playing exercises and case studies.

Everyone has their own story of poor management. The major advantage of the workshop we attended was that it was away from our home university, so that we could discuss sensitive personal situations in confidence. Some of the toughest problems are those that you might not feel comfortable about discussing with your principal investigator, your mentor or your chair.

We practised the difficult issues — how to manage meetings, for example, from distributing the agenda in advance and keeping everyone on task, to ending on a note of consensus. And through role plays, we learned how to structure negotiations as a problem-solving process rather than a battle of wills.

Except in cases of misconduct, criticism need not be personal, particularly when one is trying to motivate students. Being honest does not mean that one need be brusque or unsympathetic; we can preserve scientific integrity and encourage trainees positively.

I would strongly recommend such training. And although it is useful for postdocs, it is more crucial for young faculty members. The workshop was appealing because it was tailored to our situations by people familiar with both the academic domain and the biotech

world, where such training is more common.

Academic institutions must recognize the value of this pioneering effort and support or create such programmes for their own faculty members. They make multimillion-dollar investments in us, and, to protect their interests, should invest as seriously in leadership skills as in the progress of science.

I am already using what I learned. When I notice that I am dominating group discussions, for instance, I try to be more patient and to allow others to consider and voice their opinions. I like to think that, as a result, quieter members of my lab are becoming more confident, and that we all benefit from increased intellectual exchange. My husband has put the ideas into practice too: we wrote this article together, but were then told we could put only one name on it. Luckily, the workshop covered how to resolve authorship disputes. ■

Jessica C. Seeliger is an assistant professor at Stony Brook University School of Medicine in New York.
e-mail: jcs@pharm.stonybrook.edu

THROUGH ROLE
PLAYS, WE LEARNED
HOW TO STRUCTURE
NEGOTIATIONS
AS A PROBLEM-SOLVING
PROCESS RATHER
THAN A
BATTLE OF WILLS.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/o5xun2

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

CANCER

Blocking tumour sugar metabolism

An emerging strategy in cancer drug development is to target key metabolic molecules in tumours. Researchers have pinpointed one for prostate cancer: an enzyme involved in glucose metabolism that seems to be crucial to cancer survival.

Almut Schulze at the Cancer Research UK London Research Institute and her colleagues found that the survival of three different prostate cancer cell lines depended on glucose. Using small RNA molecules to silence genes for 222 enzymes and other molecules involved in glucose metabolism, the authors screened these cells for genes required for survival, and homed in on one, *PFKFB4*. Shutting this gene down in tumour cells stopped them from growing when they were injected into mice.

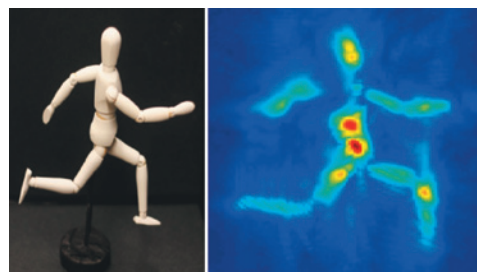
PFKFB4 enables cancer cells to produce antioxidants, which neutralize harmful oxidizing molecules. The researchers say that this protein could be a target for cancer drugs.

Cancer Discov. <http://dx.doi.org/10.1158/2159-8290.CD-11-0234> (2012)

OPTICS

Camera sees hidden objects

An ultrafast camera can create images of objects hidden behind a wall by capturing



scattered laser light.

Ramesh Raskar at the Massachusetts Institute of Technology in Cambridge and his group fired a pulse of laser light at a wall on the far side of a hidden object (**pictured, left**), and recorded the time at which the scattered light — including the small fraction of photons that bounced off the object — reached their camera. The device records images every 2 picoseconds, allowing it to record the distance travelled by each photon with sub-millimetre

precision. The team's algorithm then uses this information to reconstruct the image (**right**).

This ability to see around corners could be invaluable in dangerous or inaccessible locations, such as in highly contaminated areas or inside machinery with moving parts. *Nature Commun.* 3, 745 (2012) For a longer story on this research, see go.nature.com/nlsom5

IMMUNOLOGY

Early exposure to microbes is key

An observed increase in the prevalence of certain autoimmune diseases has been

linked to the lack of childhood exposure to microbes. A study by Dennis Kasper and Richard Blumberg at Harvard Medical School in Boston, Massachusetts, and their colleagues reveals a possible cellular mechanism for this 'hygiene hypothesis'.

The authors found that when they induced asthma or colitis in juvenile mice raised in a sterile environment, the animals had higher-than-normal levels of a type of immune cell called invariant natural killer T cells in their lungs or colon, respectively. These cells trigger inflammation and have been linked to ulcerative colitis and asthma. Moreover, expression



GEOSCIENCE

Venice: sliding down, tilting east

Although previous research had indicated that Venice had stabilized, an up-to-date study suggests that the city is still sinking — and even tilting slightly to the east.

Yehuda Bock at the University of California, San Diego, and his colleagues combined Global Positioning System data from five stations in Venice and its lagoon from 2001 to 2011 with four years of data from space-based radar

instruments. They found that Venice is sinking at a rate of 1–2 millimetres per year, with a general eastward tilt, and say that shifting tectonic plates and sediment compaction might be responsible.

The results may help the city to prepare for flooding caused by rising sea levels and seasonal tides.

Geochem. Geophys. Geosyst. <http://dx.doi.org/10.1029/2011GC003976> (2012)

of CXCL16, an inflammatory signalling molecule linked to the T cells, was also elevated in the lungs and colon, and seemed to be regulated by microbes.

Mice exposed to microbes as neonates, but not as adults, showed a decreased accumulation of the T cells, emphasizing the importance of early exposure.

Science <http://dx.doi.org/10.1126/science.1219328> (2012)

For a longer story on this research, see go.nature.com/hacoqo

REGENERATIVE BIOLOGY

Cell transplants repair colon

Tissue derived from gut stem cells can repair intestinal damage when transplanted into mice.

Mamoru Watanabe at the Tokyo Medical and Dental University, Hans Clevers at the Hubrecht Institute and University Medical Centre in Utrecht, the Netherlands, and their colleagues cultured intestinal fragments from mice and transplanted the cells, which included colonic stem cells, into mice with acute colitis. These mice gained more weight than untreated mice during the first week after treatment, and four weeks after transplantation the repaired intestinal lining seemed to be identical to the surrounding native tissue.

Colonic tissue grown from a single stem cell and placed in the mouse gut also regenerated the lining. Culturing colonic tissue from stem cells could be a therapeutic approach for human intestinal disorders such as colitis, the authors say. *Nature Med.* <http://dx.doi.org/10.1038/nm.2695> (2012)

PLANETARY SCIENCE

What lies beneath Mercury's surface

After its first year in orbit around Mercury, NASA's MESSENGER spacecraft has yielded data on the planet's

structure: the iron core is larger than previously thought and, unusually, is encased in a relatively thin shell of iron sulphide.

Maria Zuber at the Massachusetts Institute of Technology in Cambridge and her colleagues built a gravity model for the planet using measurements of tiny changes in the spacecraft's orbit. Combining this model with data on the planet's topography and spin, the authors found that as much as 85% of Mercury's radius is taken up by its dense iron core. This, along with the iron sulphide shell, helps to explain the planet's gravity field.

Another paper from Zuber and colleagues suggests that volcanic and tectonic activity persisted well past Mercury's first several hundred million years. This could explain surface features observed by the team, such as uplifted or tilted basin floors.

Science <http://dx.doi.org/10.1126/science.1218809>; <http://dx.doi.org/10.1126/science.1218805> (2012)

For a longer story on this research, see go.nature.com/orsegg

METABOLISM

Gain neurons, gain weight

Mice consuming a high-fat diet generate new neurons in a part of the brain that controls feeding and metabolism. These cells may, in turn, promote the accumulation of fat.

Seth Blackshaw at Johns Hopkins University in Baltimore, Maryland, and his colleagues found a region of brain-cell production in the hypothalamus — which regulates eating and energy use — in young adult mice. Animals fed a high-fat diet had four times the rate of neuronal production in this region, called the median eminence, than those on a normal diet.

When this brain-cell generation was blocked, mice on the fatty diet gained less weight and exhibited a

COMMUNITY CHOICE

The most viewed papers in science

GENETICS

Tracking Taz's transmissible cancer

HIGHLY READ
on www.cell.com
20 Feb–21 Mar

The contagious facial cancer devastating populations of the endangered Tasmanian devil in Australia probably originated from a female animal, a genomic analysis finds.

Elizabeth Murchison and Michael Stratton at the Wellcome Trust Sanger Institute in Hinxton, UK, and their colleagues sequenced the genomes of two healthy Tasmanian devils and two geographically distinct tumours derived from the cancer, which is spread through biting. They also analysed the genomes of 104 other tumours from across Tasmania and found that the original tumour has evolved into different subclones. Six devils had tumours with two different genetic profiles, suggesting that exposure to the cancer does not protect the animals against future bites.

Cell 148, 780–791 (2012)



D. TAL/ALBATROSS/ALAMY

speedier metabolism than animals that ate the same diet and continued to produce new neurons.

Nature Neurosci. <http://dx.doi.org/10.1038/nn.3079> (2012)

GEOPHYSICS

The reawakening of Santorini

After 60 years of silence, the volcano that erupted to form the Greek islands of Santorini (pictured) thousands of years ago seems to have reawakened.

Andrew Newman at the Georgia Institute of Technology in Atlanta and his colleagues analysed data from 24 Global Positioning System stations around the volcano from 2006 to 2012. They found that, since the beginning of 2011, the volcano's main caldera, a crater-like depression, has

been expanding by up to 18 centimetres in diameter per year — probably as a result of the expansion of its source of magma, some four kilometres below the surface. The ground deformation coincided with observations of renewed seismic activity in the area.

The earthquake activity and ground deformation could be a prelude to a small eruption, the researchers say, but a mega-eruption is unlikely. However, other volcanoes of the same type that have shown similar signs of unrest have returned to normal activity without erupting at all.

Geophys. Res. Lett. <http://dx.doi.org/10.1029/2012GL051286> (2012)

► **NATURE.COM**

For the latest research published by Nature visit:

www.nature.com/latestresearch

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

CANCER

Blocking tumour sugar metabolism

An emerging strategy in cancer drug development is to target key metabolic molecules in tumours. Researchers have pinpointed one for prostate cancer: an enzyme involved in glucose metabolism that seems to be crucial to cancer survival.

Almut Schulze at the Cancer Research UK London Research Institute and her colleagues found that the survival of three different prostate cancer cell lines depended on glucose. Using small RNA molecules to silence genes for 222 enzymes and other molecules involved in glucose metabolism, the authors screened these cells for genes required for survival, and homed in on one, *PFKFB4*. Shutting this gene down in tumour cells stopped them from growing when they were injected into mice.

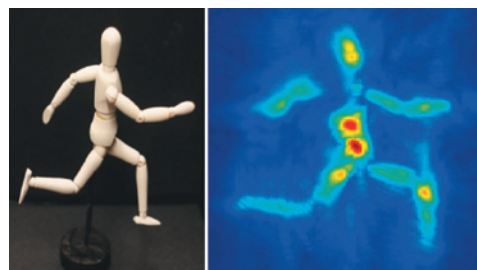
PFKFB4 enables cancer cells to produce antioxidants, which neutralize harmful oxidizing molecules. The researchers say that this protein could be a target for cancer drugs.

Cancer Discov. <http://dx.doi.org/10.1158/2159-8290.CD-11-0234> (2012)

OPTICS

Camera sees hidden objects

An ultrafast camera can create images of objects hidden behind a wall by capturing



scattered laser light.

Ramesh Raskar at the Massachusetts Institute of Technology in Cambridge and his group fired a pulse of laser light at a wall on the far side of a hidden object (**pictured, left**), and recorded the time at which the scattered light — including the small fraction of photons that bounced off the object — reached their camera. The device records images every 2 picoseconds, allowing it to record the distance travelled by each photon with sub-millimetre

precision. The team's algorithm then uses this information to reconstruct the image (**right**).

This ability to see around corners could be invaluable in dangerous or inaccessible locations, such as in highly contaminated areas or inside machinery with moving parts. *Nature Commun.* **3**, 745 (2012) For a longer story on this research, see go.nature.com/nlsom5

IMMUNOLOGY

Early exposure to microbes is key

An observed increase in the prevalence of certain autoimmune diseases has been

linked to the lack of childhood exposure to microbes. A study by Dennis Kasper and Richard Blumberg at Harvard Medical School in Boston, Massachusetts, and their colleagues reveals a possible cellular mechanism for this 'hygiene hypothesis'.

The authors found that when they induced asthma or colitis in juvenile mice raised in a sterile environment, the animals had higher-than-normal levels of a type of immune cell called invariant natural killer T cells in their lungs or colon, respectively. These cells trigger inflammation and have been linked to ulcerative colitis and asthma. Moreover, expression



GEOSCIENCE

Venice: sliding down, tilting east

Although previous research had indicated that Venice had stabilized, an up-to-date study suggests that the city is still sinking — and even tilting slightly to the east.

Yehuda Bock at the University of California, San Diego, and his colleagues combined Global Positioning System data from five stations in Venice and its lagoon from 2001 to 2011 with four years of data from space-based radar

instruments. They found that Venice is sinking at a rate of 1–2 millimetres per year, with a general eastward tilt, and say that shifting tectonic plates and sediment compaction might be responsible.

The results may help the city to prepare for flooding caused by rising sea levels and seasonal tides.

Geochem. Geophys. Geosyst. <http://dx.doi.org/10.1029/2011GC003976> (2012)

of CXCL16, an inflammatory signalling molecule linked to the T cells, was also elevated in the lungs and colon, and seemed to be regulated by microbes.

Mice exposed to microbes as neonates, but not as adults, showed a decreased accumulation of the T cells, emphasizing the importance of early exposure.

Science <http://dx.doi.org/10.1126/science.1219328> (2012)
For a longer story on this research, see go.nature.com/hacoqo

REGENERATIVE BIOLOGY

Cell transplants repair colon

Tissue derived from gut stem cells can repair intestinal damage when transplanted into mice.

Mamoru Watanabe at the Tokyo Medical and Dental University, Hans Clevers at the Hubrecht Institute and University Medical Centre in Utrecht, the Netherlands, and their colleagues cultured intestinal fragments from mice and transplanted the cells, which included colonic stem cells, into mice with acute colitis. These mice gained more weight than untreated mice during the first week after treatment, and four weeks after transplantation the repaired intestinal lining seemed to be identical to the surrounding native tissue.

Colonic tissue grown from a single stem cell and placed in the mouse gut also regenerated the lining. Culturing colonic tissue from stem cells could be a therapeutic approach for human intestinal disorders such as colitis, the authors say. **Nature Med.** <http://dx.doi.org/10.1038/nm.2695> (2012)

PLANETARY SCIENCE

What lies beneath Mercury's surface

After its first year in orbit around Mercury, NASA's MESSENGER spacecraft has yielded data on the planet's

structure: the iron core is larger than previously thought and, unusually, is encased in a relatively thin shell of iron sulphide.

Maria Zuber at the Massachusetts Institute of Technology in Cambridge and her colleagues built a gravity model for the planet using measurements of tiny changes in the spacecraft's orbit. Combining this model with data on the planet's topography and spin, the authors found that as much as 85% of Mercury's radius is taken up by its dense iron core. This, along with the iron sulphide shell, helps to explain the planet's gravity field.

Another paper from Zuber and colleagues suggests that volcanic and tectonic activity persisted well past Mercury's first several hundred million years. This could explain surface features observed by the team, such as uplifted or tilted basin floors.

Science <http://dx.doi.org/10.1126/science.1218809>; <http://dx.doi.org/10.1126/science.1218805> (2012)
For a longer story on this research, see go.nature.com/orsegg

METABOLISM

Gain neurons, gain weight

Mice consuming a high-fat diet generate new neurons in a part of the brain that controls feeding and metabolism. These cells may, in turn, promote the accumulation of fat.

Seth Blackshaw at Johns Hopkins University in Baltimore, Maryland, and his colleagues found a region of brain-cell production in the hypothalamus — which regulates eating and energy use — in young adult mice. Animals fed a high-fat diet had four times the rate of neuronal production in this region, called the median eminence, than those on a normal diet.

When this brain-cell generation was blocked, mice on the fatty diet gained less weight and exhibited a

COMMUNITY CHOICE

The most viewed papers in science

GENETICS

Tracking Taz's transmissible cancer

HIGHLY READ
on www.cell.com
20 Feb–21 Mar

The contagious facial cancer devastating populations of the endangered Tasmanian devil in Australia probably originated from a female animal, a genomic analysis finds.

Elizabeth Murchison and Michael Stratton at the Wellcome Trust Sanger Institute in Hinxton, UK, and their colleagues sequenced the genomes of two healthy Tasmanian devils and two geographically distinct tumours derived from the cancer, which is spread through biting. They also analysed the genomes of 104 other tumours from across Tasmania and found that the original tumour has evolved into different subclones. Six devils had tumours with two different genetic profiles, suggesting that exposure to the cancer does not protect the animals against future bites.

Cell 148, 780–791 (2012)



D. TAL/ALBATROSS/ALAMY

speedier metabolism than animals that ate the same diet and continued to produce new neurons.

Nature Neurosci. <http://dx.doi.org/10.1038/nn.3079> (2012)

GEOPHYSICS

The reawakening of Santorini

After 60 years of silence, the volcano that erupted to form the Greek islands of Santorini (pictured) thousands of years ago seems to have reawakened.

Andrew Newman at the Georgia Institute of Technology in Atlanta and his colleagues analysed data from 24 Global Positioning System stations around the volcano from 2006 to 2012. They found that, since the beginning of 2011, the volcano's main caldera, a crater-like depression, has

been expanding by up to 18 centimetres in diameter per year — probably as a result of the expansion of its source of magma, some four kilometres below the surface. The ground deformation coincided with observations of renewed seismic activity in the area.

The earthquake activity and ground deformation could be a prelude to a small eruption, the researchers say, but a mega-eruption is unlikely. However, other volcanoes of the same type that have shown similar signs of unrest have returned to normal activity without erupting at all.

Geophys. Res. Lett. <http://dx.doi.org/10.1029/2012GL051286> (2012)

► **NATURE.COM**

For the latest research published by Nature visit:

www.nature.com/latestresearch

SEVEN DAYS

The news in brief

RESEARCH

Laser fusion

The US National Ignition Facility (NIF) revealed last week that its 192 lasers had combined to fire a shot with an energy of 1.875 megajoules — a milestone in efforts to trigger fusion by imploding a frozen fuel pellet of hydrogen isotopes. The energy means that the facility, at the Lawrence Livermore National Laboratory in Livermore, California, has surpassed its design specifications. But the shot was just a demonstration, and did not involve a fuel target; the NIF is still racing to achieve 'ignition' (in which more energy is got out from fusion than is put in by the laser) by the end of this year. See go.nature.com/giv6ru for more.

Einstein online

An online archive of Albert Einstein's personal papers and related documents is being expanded and updated, the Hebrew University of Jerusalem announced on 19 March. A limited sample of the physicist's papers is already available, but the digitization project, funded by the Polonsky Foundation UK (which helped to digitize Isaac Newton's archives), will see more than 80,000 documents put online. See go.nature.com/owyoxd for more.

POLICY

Lessons from Potti

Lapses in oversight that stopped a US university from halting clinical trials based on flawed research are symptomatic of a larger problem, according to a 23 March report by the Institute of Medicine in Washington DC. Failures in research oversight, data management and clinical-trial design permitted trials to proceed



M. THIESSEN/NATIONAL GEOGRAPHIC

Back from the abyss

Canadian film director James Cameron returned from his solo dive to the bottom of the 11-kilometre-deep Mariana Trench in the Pacific Ocean on 26 March, describing it as a gelatinous landscape as desolate as the Moon. Cameron (pictured, in his submersible) had hoped to collect samples of water and rock from

the apparently lifeless sediment, but technical problems got in the way, so the team is hoping for three or four more dives in coming weeks. It was the first manned visit to the trench, the ocean's deepest spot, since the *Trieste* submersible visited the region in 1960. See go.nature.com/er8ag6 for more.

even though they were based on faulty research by cancer geneticist Anil Potti, then at Duke University in Durham, North Carolina (see *Nature* **469**, 139–140; 2011). These problems could exist at other institutions, the report warned, calling for higher standards in other tests based on large-scale genomic and proteomic studies. See go.nature.com/hofgoc for more.

No Vatican meeting

The Vatican has abruptly cancelled a controversial stem-cell conference that was to have included an audience with the Pope next month. The Third International Congress on Responsible Stem Cell Research, scheduled for 25–28 April, was set to focus on clinical applications of

adult and reprogrammed stem cells, although some of the invited speakers do research using human embryonic stem cells, which the Catholic Church considers unethical. The Church's Pontifical Academy for Life, one of the conference organizers, said that the cancellation was forced by logistical, organizational and financial factors. See go.nature.com/uhapin for more.

Future Earth

An alliance of researchers, United Nations bodies and global science funders are teaming up to form a ten-year research initiative that will link environmental change with understanding of its impact on human development. The initiative, called 'Future

Earth — research for global sustainability', was announced on 27 March at the Planet Under Pressure conference in London. Officially operational from 2013, the programme will try to unite social sciences and humanities with environmental science. See go.nature.com/s14iyj for more.

Primate deaths

The New Iberia Research Center near Lafayette, Louisiana, is being investigated by government regulators for an incident last year in which three decomposing macaques were found trapped in a metal chute connecting two outdoor cages. The US Department of Agriculture (USDA) probe was publicized by animal activists at a 27 March press conference. Last week, activists

also publicized two other instances of primate deaths at research laboratories: a January USDA inspection report recording the second death of a primate in less than six months at drug firm Bristol-Myers Squibb (headquartered in New York City); and a February warning letter sent to Rockefeller University in New York City, after a macaque died last year when its collar became entangled with another animal's. See go.nature.com/r1xrqo for more.

Antibiotic ban

The US Food and Drug Administration (FDA) might issue further restrictions on the use of antibiotics in farm animals, after a federal court in New York ordered it to take action on a 1977 finding that dosing livestock with penicillin and tetracyclines could promote the spread of drug-resistant bacteria. The FDA started proceedings to restrict the drugs but never completed them — and in December 2011 formally abandoned the process. The ruling on 22 March, after a lawsuit by a coalition of watchdog groups, means that the agency must continue where it left off.

Neutrino worry

The future of a pioneering US project to study neutrinos was thrown into doubt on 19 March, when officials

at the US Department of Energy (DOE) said that they were reluctant to fund it in its current form. Expected to come online in 2022–24, the Long-Baseline Neutrino Experiment would use more than 30,000 tonnes of liquid argon housed in the Homestake Mine near Lead, South Dakota, to detect neutrinos sent 1,300 kilometres from Fermilab in Batavia, Illinois. But it would cost between US\$1.2 billion and \$1.5 billion — eating up too much of the DOE's budget for high-energy physics. See go.nature.com/awptgm for more.

PEOPLE

Abel award

Mathematician Endre Szemerédi, of the Alfréd Rényi Institute of Mathematics in Budapest, has been awarded this year's Abel Prize in mathematics, worth around US\$1 million and considered to be as prestigious as the Nobel prize. Szemerédi won for his work on discrete mathematics (relating to discrete entities such as number sequences, logic operations, and networks) and theoretical computer science. See go.nature.com/evizub for more.

World Bank leader

Jim Yong Kim, an expert in the field of global health, has been nominated by US President



Barack Obama to be the next leader of the World Bank. Kim, currently president of Dartmouth College in Hanover, New Hampshire, is a physician and anthropologist who led the World Health Organization's HIV/AIDS unit during 2004–06. As the US choice, Kim (pictured) is almost assured of the post, although his 23 March nomination to replace Robert Zoellick was unexpected: it would be the first time that the World Bank was not headed by a financier, economist or politician. See go.nature.com/xnzw8c for more.

Gairdner awards

Jeffrey Ravetch, from the Rockefeller University in New York City, is among seven scientists who have won this year's Canada Gairdner Awards. The prizes, each worth US\$100,000, are given by the Gairdner Foundation in Toronto for leading biomedical research and often presage Nobel prizes. Ravetch

COMING UP

31 MARCH–4 APRIL

The American Association for Cancer Research meets in Chicago, Illinois, with recent studies on cancer genetics to the fore. go.nature.com/a76fdh

3–4 APRIL

The Royal Society in London hosts a meeting to discuss flu-virus research — including concerns about biosecurity. go.nature.com/9ngnx7

won for his work on immune receptors. Other winners announced on 22 March include Michael Young (also at Rockefeller), Jeffrey Hall and Michael Rosbash (both at Brandeis University in Waltham, Massachusetts) for their work on biological clocks. See go.nature.com/z4zoe8 for more.

BUSINESS

Patent panic

The US Supreme Court has ordered a New York appeals court to reconsider its ruling last year that patents on genes are valid (see *Nature* 476, 11; 2011). The 26 March order follows a 20 March Supreme Court ruling to overturn two patents on a way to determine drug dosage because they were based on the laws of nature — a decision that rocked the biotech industry. Those patents, owned by the biotech company Prometheus Laboratories in San Diego, California, covered the process of administering a class of drug called thiopurines and measuring blood levels of key metabolites to determine whether the dose received is safe and effective. See go.nature.com/lj9kyl for more.

➤ **NATURE.COM**

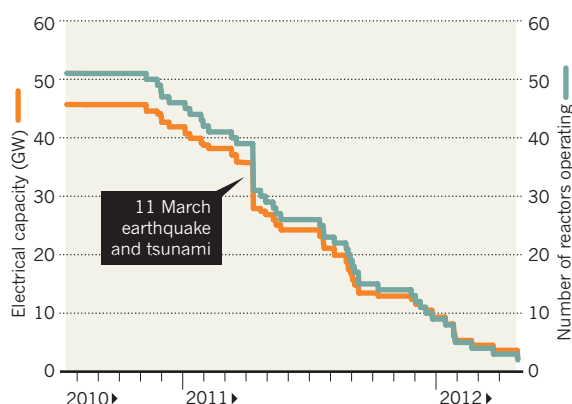
For daily news updates see: www.nature.com/news

TREND WATCH

Only one nuclear reactor is currently operating in Japan (see chart), after the Tokyo Electric Power Company took a reactor at the Kashiwazaki-Kariwa power station in Niigata Prefecture offline for routine maintenance on 26 March. None of the reactors closed after an earthquake and tsunami struck the Fukushima Daiichi plant last March has yet reopened, although some have passed official 'stress tests' showing that they could withstand similar disasters.

JAPAN'S NUCLEAR SHUTDOWN

Only one nuclear reactor is now operating after last year's earthquake and tsunami.



NEWS IN FOCUS

FLU *Nature* analysis highlights holes in global virus surveillance **p.520**

EXOPLANETS Weighing the Kepler candidates with Swiss precision **p.522**

CHEMISTRY Millions of molecules set free in public databases **p.524**

AGRICULTURE Ways to treat Africa's ailing soils **p.525**



L. DE ALMEIDA/CONTRASTO/EYEVINE



Soya growers in Brazil who sign up to proposed environmental standards would be able to increase the productivity of existing farmland and so limit deforestation.

ENVIRONMENT

Farm focus for saving trees

Round-table talks aim to slow climate warming by transforming agriculture.

BY JEFF TOLLEFSON

The principle is seductively simple: to reduce carbon emissions, leave tropical forests standing. But a widely heralded approach in which rich nations would pay poorer ones to keep their forests intact has proved trickier to deploy than many had hoped. Now a consortium of scientists, environmentalists and industries is expanding the focus from preserving forests to tackling the main driver of deforestation: agriculture.

The United Nations forestry initiative — known as REDD, for Reducing Emissions from Deforestation and Forest Degradation — was originally seen as a way of changing frontier economics by attaching a monetary

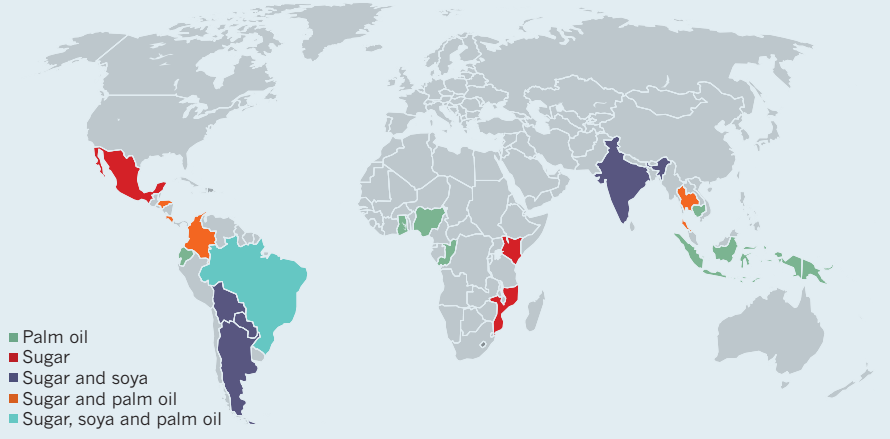
value to standing forests, which take up carbon dioxide and stabilize the climate. Carbon payments would make it easier for landowners to earn a living without clearing more land. But despite years of negotiations and several billion dollars in commitments, little money has filtered down to those who live and work at the forest frontier. Where money has changed hands, it has happened mostly among governments, says Daniel Nepstad, a US ecologist who heads the international programmes for the Amazon Environmental Research Institute (IPAM), headquartered in Brasilia. As a result, he says, scepticism is rising among those who are supposed to benefit most.

Nepstad and others involved in the latest REDD effort see potential for faster progress

by merging REDD initiatives with a series of 'commodity round tables', which bring multinational companies such as beverage firm PepsiCo, agricultural biotech giant Monsanto and retailer Walmart together with producers and environmentalists to negotiate environmental certification standards for products such as soya beans, palm oil, sugar cane and beef. These standards focus on everything from soil management to workers' rights, and include limits on deforestation. The idea is that producers who sign up and implement best practices will be able to increase productivity, command a higher price for their products and pressure competitors to raise their own standards. "REDD has become loaded down," Nepstad says. "Where it is moving forward ►

FOOD VERSUS FORESTS

More sustainable farming of key crops could slow the rate of tropical deforestation. The map identifies countries and crops that offer the greatest gains.



► most effectively is where it is moving together with rural development strategies.”

IPAM leads a loose-knit consortium known as the Roundtable-REDD, which released an analysis on 28 March that identifies countries in which investing in projects for production of sugar cane, soya and palm oil could have the greatest impact on carbon (see ‘Food versus forests’). With more than US\$4 million in seed money from Norway, the consortium plans to announce an initial round of projects in the run-up to the UN Conference on Sustainable Development in Rio de Janeiro (Rio+20) in Brazil in June.

Initially, consortium projects could focus on restoring abandoned agricultural lands and intensifying production on existing farms, ranches and plantations. This could involve accelerated replanting of palm plantations with high-yield varieties in Indonesia, or helping farmers and ranchers to access existing money for sustainable agriculture in Brazil. In the latter case, Nepstad says that the consortium is looking at ways to help those landowners who commit to round-table standards apply for subsidized government loans — worth some \$1.7 billion during the current growing season — so that they can improve soils and intensify their production or redevelop degraded fields instead of clearing new ones. The consortium is also looking at tackling overall greenhouse-gas emissions using carbon credits, which could be sold to

private investors or companies seeking to offset their own emissions. Those funds could help farmers and ranchers bring in better soil-management practices, make more efficient use of fertilizers or capture methane emissions for electricity generation.

GAME CHANGER

The programme’s potential is highlighted by recent progress in Brazil, where Amazon deforestation has declined by 78% since a 2004 peak even as agricultural production continued to climb. Recent studies suggest that government enforcement and broader agricultural policies^{1,2} have played a part, but consumers and environmentalists have also contributed by pressuring major food suppliers to sign moratoria on the purchase of soya and beef from recently cleared land. The round-table model, which is already operating for some commodities, is similar. Although it is too early to see land-use changes in satellite data, the round tables do seem to be affecting the way many companies do business, says Holly Gibbs, an environmental geographer at the University of Wisconsin-Madison. “I don’t know that it’s a sea change yet,” Gibbs says, “but they are definitely changing the rules and the norms and the way these industries operate.”

Nonetheless, even advocates acknowledge that it is difficult to achieve consensus on environmental standards in a room of producers, major food companies and environmentalists,

all with competing agendas. This has led to some criticism that round tables provide political cover for companies that wish to avoid making stronger commitments to the environment.

“It’s my contention that the round tables are holding back innovation,” says Scott Poynton, executive director of the Forest Trust, a non-profit organization in Crassier, Switzerland, which has been working with Swiss food company Nestlé and the world’s second-largest palm-oil producer, Golden Agri-Resources in Indonesia, to enforce their zero-deforestation commitments. Companies that want to halt deforestation should pressure their suppliers directly, Poynton says. “The model is there,” he says, “and it doesn’t require the vast billions of dollars everybody is talking about.”

Instead, the round tables are intended to help propagate minimum environmental standards across the entire world. “It takes a while before you have all of the companies aligned on these principles,” says Jeroen Douglas, South American programme director for the Solidaridad network in Buenos Aires, which focuses on sustainable supply chains. Solidaridad plans to invest around €70 million (US\$94 million) to help some 400,000 small-scale farmers and ranchers to achieve round-table certification by 2015. But there are barriers, and Douglas says that the link with REDD money might be enough to seduce local producers.

Other efforts are also emerging. For example, a coalition of state-level governments in the United States, Brazil, Indonesia, Nigeria, Peru and Mexico has launched its own initiative, the Governors’ Climate and Forests Task Force. It is working to set up mechanisms that would allow companies in participating regions to offset emissions by paying to reduce deforestation.

William Boyd, a law professor at the University of Colorado in Boulder and project leader for the task force, says that both the round-table and state efforts are emblematic of what is needed. If the money to transform agriculture and reduce the incentives for clearing forests doesn’t begin to flow soon, farmers in the developing world will give up on the process, Boyd says. “And who could blame them?” ■

1. Rudorff, B. F. T. *et al. Remote Sens.* **3**, 185–202 (2011).
2. Macedo, M. N. *et al. Proc. Natl Acad. Sci. USA* **109**, 1341–1346 (2012).

MORE ONLINE

TOP STORY

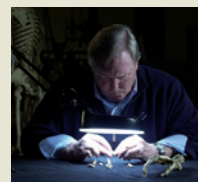


Titanium signature poses puzzle for popular theory of Moon’s origin
go.nature.com/bp1amg

MORE NEWS

- Microbiome sequencing offers hope for diagnostics go.nature.com/bownno
- Chinese data confirm that more girls than boys are born in famine go.nature.com/pfqlbn
- Earthquake tests 25 years of Mexican engineering go.nature.com/s6utpp

VIDEO



Fossil footprints hint at different ways for early humans to walk
go.nature.com/onrjrh

CANCER

Drug candidates derailed in case of mistaken identity

PARP inhibitor that wasn't highlights widespread flaws in preclinical studies.

BY HEIDI LEDFORD

Scott Kaufmann fears a rush to judgement. He studies DNA-repair proteins called poly(ADP-ribose) polymerases, or PARPs, that have shown promise as targets for anticancer drugs. Until early last year, compounds that inhibit PARPs were the next big thing in cancer drug development (see 'Suspect class'). Then, a leading candidate PARP inhibitor called iniparib failed a phase III clinical trial for a form of breast cancer, and Kaufmann, a biomedical scientist at the Mayo Clinic in Rochester, Minnesota, was dismayed to find that cancer researchers seemed to be giving up on PARP inhibitors as a whole. "I got tired of hearing from clinical colleagues that iniparib failed, therefore PARPs are a terrible target," he says. "Clinicians were saying they didn't want to open any more clinical trials of PARP inhibitors."

Kaufmann decided to take a closer look at iniparib, which was developed by Paris-based pharmaceutical company Sanofi. PARP inhibitors effectively make cells more susceptible to DNA damage. Cancer cells already hit by DNA-damaging chemotherapies, or tumours bearing mutations that inactivate DNA-repair pathways, are particularly sensitive to the drugs. Kaufmann and his collaborators tested iniparib on cancer cells grown in the lab, looking for signs of PARP inhibition. They found none. Their results, published earlier this year alongside work from other groups^{1,2}, add to a growing body of evidence that iniparib may not be a potent PARP inhibitor after all.

So why did preclinical tests, which map out a drug's mechanism of action before it goes into human trials, get it so wrong? Most of the early studies on iniparib are unpublished, leaving researchers guessing what might have gone amiss. But those in the field knew that the drug seemed to lack potency: lab assays required tremendous concentrations of the compound to show any effect on PARP proteins.

Yet Kaufmann and others say that researchers are still misinterpreting the drug's failure, casting doubt on what could yet be promising targets. "It besmirched the entire class of compounds," says Alan Ashworth, chief executive of the Institute of Cancer Research in London, who has developed other PARP inhibitors.

Unreliable results from preclinical studies are nothing new to the pharmaceutical

SUSPECT CLASS

Despite flagging enthusiasm, several PARP-inhibitor drugs are still making their way through clinical trials.

Drug	Company	Stage	Cancers
Olaparib (AZD-2281)	AstraZeneca	Phase I	Solid tumours
Veliparib (ABT-888)	Abbott	Phase I, II	Prostate, colorectal, leukaemias, solid tumours
CEP-9722	Cephalon	Phase I, II	Solid tumours, lymphoma
Rucaparib (CO-338)	Clovis Oncology	Phase I, II	BRCA1/2 mutant cancers, solid tumours
E7016	Eisai	Phase I	Solid tumours
BMN-673	BioMarin Pharmaceutical	Phase I	Leukaemias, solid tumours

Source: ClinicalTrials.gov, data accessed 26 March 2012.

industry. But the problem is acquiring greater urgency as drug companies look for new ways to lower failure rates and trim budgets. Between 2008 and 2009, only 18% of drugs in phase II clinical trials succeeded³. And, as described in a Comment in this issue (see page 531), when the biotechnology company Amgen, based in Thousand Oaks, California, tried to reproduce data from 53 published preclinical studies of potential anticancer drugs, it failed in all but six cases.

The reproducibility problem isn't limited to published studies. Comment co-author Glenn Begley, former head of haematology and oncology at Amgen and now a freelance consultant, told *Nature* that Amgen had evaluated hundreds of potential projects from biotechnology firms each year, with an eye to selecting new partnerships. Many of those

"Probably there was too much hype in the first place. Now it's too negative. Perhaps it will balance out."

projects were also irreproducible, he says. Although unintentional bias rather than fraud accounts for most of the false leads, they still waste desperately needed resources.

"At the end of this there is a patient," Begley says, noting that several of the studies that his team at Amgen could not reproduce had already spawned clinical trials. "It's a distraction, and the drug-development challenge before us is already so great."

In the case of iniparib, early success in a small phase II clinical trial was quickly overtaken by negative results in a larger study, which Sanofi announced in January 2011. More bad news for PARP inhibitors followed

in December, when AstraZeneca, a London-based pharmaceutical firm, revealed that its leading PARP inhibitor, olaparib, had not performed well enough in a phase II clinical trial against ovarian cancer to warrant continued investment.

There is no doubt in the field that olaparib is a bona fide PARP inhibitor, says Susan Domchek, an oncologist at the University of Pennsylvania in Philadelphia. She and others suspect that olaparib did poorly because it was tested in a broad population of cancer patients, rather than being targeted to those most likely to benefit.

For example, cancer-promoting mutations in the breast-cancer genes *BRCA1* and *BRCA2* also disable DNA-repair pathways, and studies have shown that patients with these mutations do respond to PARP inhibitors⁴. Clinicians were dismayed when AstraZeneca told them that plans to test olaparib specifically in patients who carry the *BRCA* mutations were also on hold following the ovarian cancer results. Nevertheless, "after these failures, people are again looking more carefully at this population of *BRCA1* and *BRCA2* mutation carriers", says Domchek.

In the end, some dampening of enthusiasm for PARP inhibitors may be healthy, says Ashworth. "Probably there was too much hype in the first place," he concedes. "Now it's too negative. Eventually, perhaps it will balance out." ■ SEE EDITORIAL P.509

1. Patel, A. G., De Lorenzo, S. B., Flatten, K. S., Poirier, G. G. & Kaufmann, S. H. *Clin. Cancer Res.* **18**, 1655–1662 (2012).
2. Liu, X. *et al. Clin. Cancer Res.* **18**, 510–523 (2012).
3. Arrowsmith, J. *Nature Rev. Drug Discovery* **10**, 328–329 (2011).
4. Tutt, A. *et al. Lancet* **376**, 235–244 (2010).

Flu surveillance lacking

Nature analysis highlights need for international strategy to watch for pandemic threats.

BY DECLAN BUTLER

When researchers created strains of the H5N1 avian influenza virus that could spread easily between mammals, they argued that their work would aid in surveillance, by identifying mutations to watch for in the wild.

But an analysis by *Nature* paints a dire picture of how animal flu viruses are being monitored. In 2010, the world's poultry population was estimated at 21 billion, yet only around 1,000 flu sequences from 400 avian virus isolates were collected — and many countries that are home to billions of farmed chickens, ducks and pigs contributed few or none.

In addition, the surveillance is typically not sustained, but instead is ad hoc and reactive, and is largely in response to disease outbreaks or temporary research projects. But a flu virus that emerges anywhere, at any time, can threaten the entire planet. The *Nature* analysis “highlights a global problem: lack of data”, says Ian Brown, head of avian virology and mammalian influenza at the Animal Health and Veterinary Laboratories Agency lab in Weybridge, UK.

Timely global surveillance of animal flu viruses is crucial not just for identifying pandemic threats, but also for detecting outbreaks, monitoring how viruses are evolving, understanding risk factors that enable them to spread and keeping animal vaccines and diagnostics up to date.

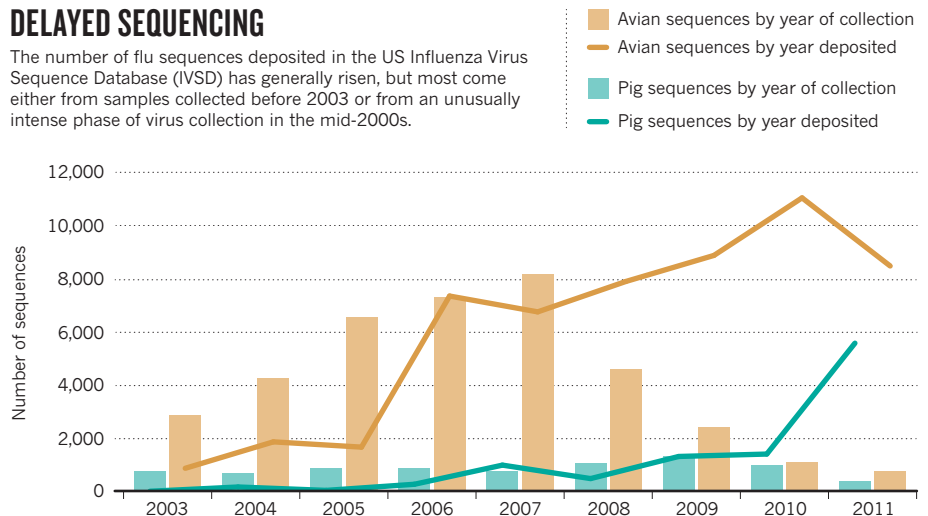
To assess trends in global genetic surveillance, *Nature* analysed the records of non-identical sequences from all subtypes of avian and pig flu deposited in the US National Center for Biotechnology Information's Influenza Virus Sequence Database between 2003 and 2011. The database contains sequences from GenBank and several large flu sequencing projects, including the Influenza Genome Sequencing Project — a major initiative run by the National Institute of Allergy and Infectious Diseases (NIAID) to boost the sequencing of existing isolates. The analysis covered all subtypes of flu virus, not just H5N1. That's important, says Malik Peiris, a flu virologist and surveillance expert at the University of Hong Kong, because “H5N1 is not the sole pandemic candidate, and low pathogenic viruses are just as likely, if not more likely, to become pandemic”.

The number of avian flu sequences deposited in the database skyrocketed between 2003 and

➔ NATURE.COM
For more see
Nature's mutant flu
special:
go.nature.com/mhmibi

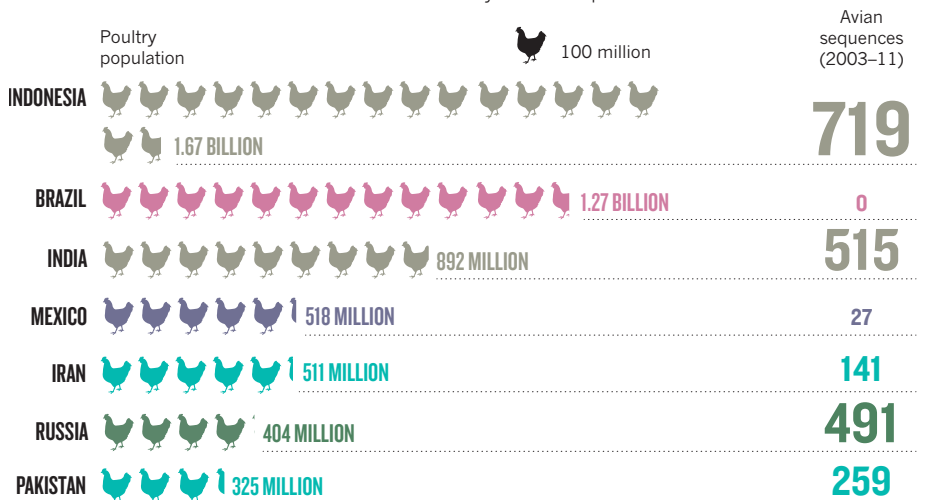
DELAYED SEQUENCING

The number of flu sequences deposited in the US Influenza Virus Sequence Database (IVSD) has generally risen, but most come either from samples collected before 2003 or from an unusually intense phase of virus collection in the mid-2000s.



MANY BIRDS, FEW SAMPLES

Some countries with large poultry populations have generated only a small number of sequences, either because outbreaks of avian flu are rare there or because they have inadequate surveillance.



2010, before dropping off in 2011. The number of pig sequences deposited remained relatively flat from 2003 to 2010, before jumping dramatically in 2011.

However, few contemporary data are available. The number of avian flu sequences from isolates collected in each year peaks in 2007 and plummets thereafter. The jump in the number of pig sequences also disappears (see ‘Delayed sequencing’).

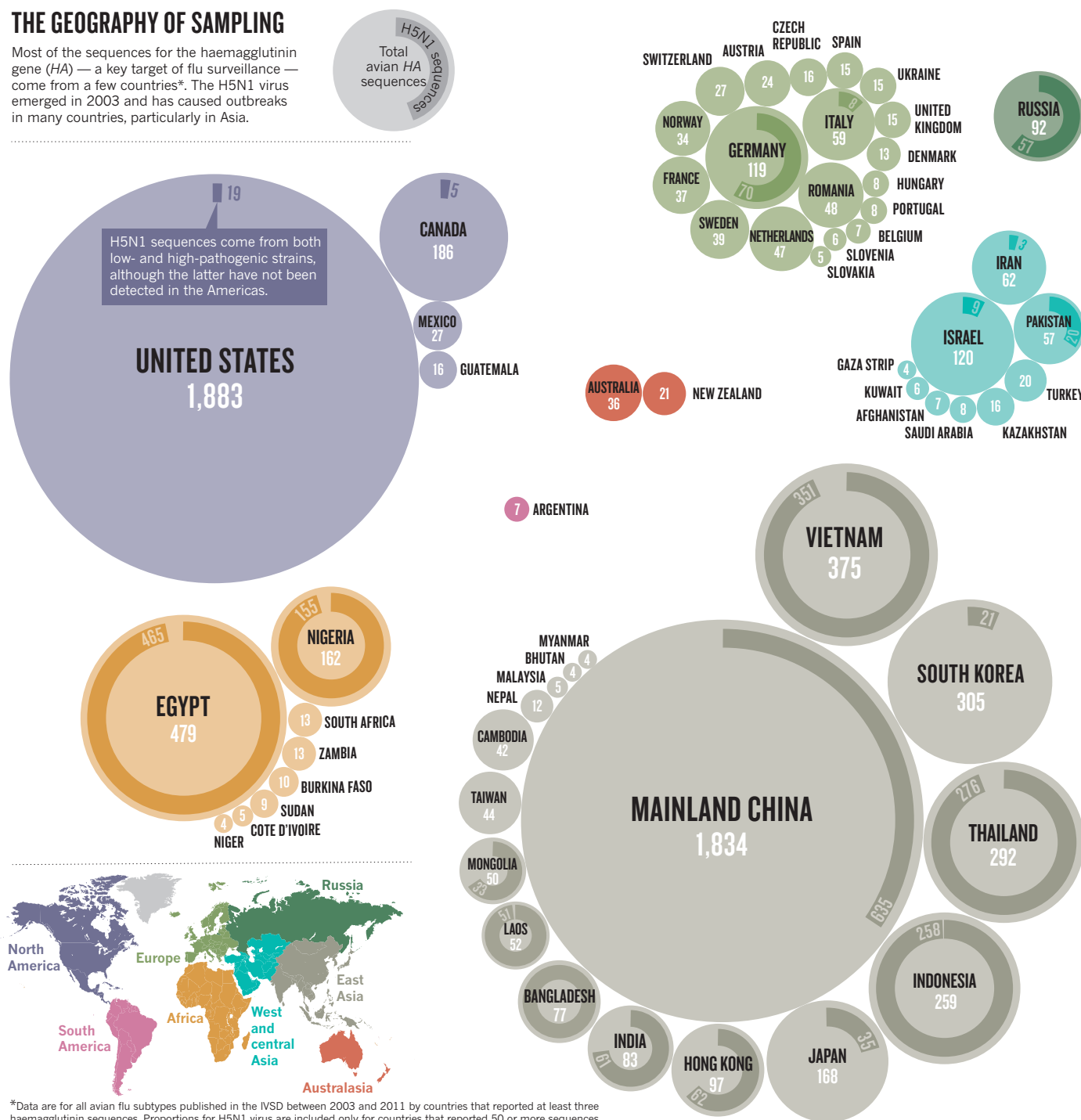
Roughly 30% of the sequences are from isolates collected before 2003. The 2007 peak in avian viral sampling was largely the result of surveys of more than 100,000 wild birds to

monitor for the arrival of H5N1 in the Americas^{1,2}. Also contributing was the sequencing of the H5N1 viral flare that moved from Asia into Europe and Africa in 2005 and 2006 (refs 3, 4). The older sequences can inform surveillance by showing how the viruses have evolved, says Peiris, but contemporary data are important “for real-time surveillance”, such as spotting changes that might herald dangerous strains.

Many years can pass between the collection and sequencing of isolates, says Sylvie van der Werf, head of the Molecular Genetics of RNA Viruses lab at the Pasteur Institute in Paris. One reason is that many of the virus samples

THE GEOGRAPHY OF SAMPLING

Most of the sequences for the haemagglutinin gene (HA) — a key target of flu surveillance — come from a few countries*. The H5N1 virus emerged in 2003 and has caused outbreaks in many countries, particularly in Asia.



*Data are for all avian flu subtypes published in the IVSD between 2003 and 2011 by countries that reported at least three haemagglutinin sequences. Proportions for H5N1 virus are included only for countries that reported 50 or more sequences.

are sequenced in retrospective research studies. Another is a lack of funding for sequencing, although falling sequencing costs are easing this bottleneck. The Influenza Genome Sequencing Project is also helping by generating vast quantities of sequences — it now accounts for half of all avian and pig sequences — but it is aimed at increasing the genomic knowledge base, rather than real-time surveillance.

Researchers, too, are contributing to the lag, because many do not share their sequences until after the data have been published. An exception is the Centers of Excellence for Influenza Research and Surveillance — a network created

by the NIAID in 2007 to boost flu surveillance — which has a policy of releasing all sequence data within 45 days of its collection.

The two agencies responsible for monitoring disease outbreaks in animals — the Food and Agriculture Organization (FAO) of the United Nations and the World Organisation for Animal Health (OIE) — stipulate that sequences of potentially zoonotic viruses should be deposited in public databases within 3 months, but few researchers do so, says Ilaria Capua an avian-flu researcher at the Veterinary Public Health Institute in Legnaro, Italy, who champions greater availability of sequences⁵.

Nature also looked at where the sequences come from (see 'The geography of sampling'). The picture that emerges is worse than some experts had thought. Almost all come from just a handful of countries — most countries have little or no genetic surveillance in place.

Just 7 of the 39 countries with more than 100 million poultry in 2010 collected more than 1,000 avian flu samples between 2003 and 2011. Eight countries — Brazil, Morocco, the Philippines, Colombia, Ecuador, Algeria, Venezuela and the Dominican Republic — collected none at all; 13 collected between 1 and 100; and 11 collected between 100 and ▶

► 1,000. Even fewer pig sequences were collected, with one-third of the countries that are home to more than 4 million pigs depositing none at all.

The size of a country's poultry population is no predictor of how many samples that country will generate (see 'Many birds, few samples'). Countries that have well-developed veterinary services and a well-structured and hygienic farming industry inevitably have fewer flu sequences to report, as disease levels tend to be low, says Brown. However, many of the countries that have contributed few or no sequences have poor veterinary systems and flu-prone farming systems, such as backyard farms and mixed poultry and pig farms, which are often close to wild ducks and other flu reservoirs.

"Proper geographic representation is lacking," says van der Werf, as is sustained surveillance. This results in large gaps in data, she says, because "many consecutive years of surveillance are needed to see trends" (see page 535). Poorer countries tend to have inadequate surveillance resources, and farmers often have little incentive to report outbreaks because they will not receive any compensation for culled livestock. Countries sometimes also fail to look for, or report, outbreaks so that they can claim they are free of infection and so avoid trade problems.

Flu experts say that the dire state of surveillance could be rapidly turned around by, for example, creating a network of sentinel sites, focusing on the countries and regions most at risk, that would collect isolates and sequence them in real time. Such a network would probably even cost less than the fragmented and uncoordinated surveillance efforts in place today, says Jeremy Farrar, director of the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam (see page 534).

The problem is that no global body has overall responsibility for flu surveillance. The World Health Organization (WHO) runs a global network of labs for human flu surveillance and selects human strains to be included in vaccines for seasonal flu. Monitoring animals falls to the FAO, which tends to focus on food security, and the OIE, which looks mostly at animal health and trade.

What is needed is international leadership, says Farrar. "If, say, the WHO and the FAO were to construct an advisory framework, surveillance could probably be done much more systematically and efficiently."

■ SEE EDITORIAL P. 509

1. Butler, D. & Ruttimann, J. *Nature* **441**, 137–139 (2006).
2. Check, E. *Nature* **442**, 348–350 (2006).
3. Butler, D. *Nature* <http://dx.doi.org/10.1038/news050801-1> (2005).
4. Butler, D. *Nature* <http://dx.doi.org/10.1038/news060206-7> (2006).
5. *Nature* **440**, 255–256 (2006).



Italy's National Galileo Telescope in the Canary Islands will host the HARPS-North planet finder.

SPACE

North set for mass analysis of planets

Spectrograph will review results from Kepler telescope.

BY ERIC HAND

How many extrasolar planets has NASA's Kepler mission discovered? That depends on how you count. Last month, the mission team published a catalogue that lists a staggering 2,321 candidate planets, amassed since May 2009 as the space-based telescope watches stars for the shadow of planets passing over their faces. Yet only 69 of them are considered confirmed planets. Astronomers have fretted over the growing backlog, but help is on its way.

For a Kepler planet to ascend from candidate to confirmed, a second method has to vouch for it: for example, a ground-based spectrograph must report signs that the planet's gravity is tugging its star back and forth. Yet Kepler looks north, whereas the instrument most sensitive to stellar wobbles, the European Southern Observatory's High Accuracy Radial Velocity Planet Searcher (HARPS), is located at the La Silla Observatory in Chile and can only observe the southern sky. On 1 April, however, the Northern Hemisphere will get a near-clone of HARPS when HARPS-North achieves first light at the Italian 3.6-metre National Galileo Telescope (TNG) on La Palma in the Canary Islands.

The instrument has been a long time coming. Conceived in 2005, the project was originally led by Harvard University in Cambridge, Massachusetts. But in 2010, after Harvard's endowment fell during the financial crisis, the University of Geneva in Switzerland took charge. Financial problems forced the group to switch from the 4.2-metre William Herschel Telescope, also on La Palma, to the TNG, which will give the HARPS-North team 80 nights of dedicated time per year for five years.

That should help to alleviate the bottleneck for Kepler candidates. Many astronomers, however, are looking to HARPS-North less for confirmation of the candidate planets than for insight into their properties. The false-positive rate for Kepler, after all, has already been shown to be less than 10% (T. D. Morton and J. A. Johnson *Astrophys. J.* (in the press) Preprint at <http://arxiv.org/abs/1101.5630>; 2011). "It has become acceptable to do a statistical analysis and say, 'They are planets,'" says Joshua Winn, an astronomer at the Massachusetts Institute of Technology in Cambridge.

What Winn and

➔ **NATURE.COM**

For more on Kepler's search for exoplanets, visit: go.nature.com/ptgbbt

► 1,000. Even fewer pig sequences were collected, with one-third of the countries that are home to more than 4 million pigs depositing none at all.

The size of a country's poultry population is no predictor of how many samples that country will generate (see 'Many birds, few samples'). Countries that have well-developed veterinary services and a well-structured and hygienic farming industry inevitably have fewer flu sequences to report, as disease levels tend to be low, says Brown. However, many of the countries that have contributed few or no sequences have poor veterinary systems and flu-prone farming systems, such as backyard farms and mixed poultry and pig farms, which are often close to wild ducks and other flu reservoirs.

"Proper geographic representation is lacking," says van der Werf, as is sustained surveillance. This results in large gaps in data, she says, because "many consecutive years of surveillance are needed to see trends" (see page 535). Poorer countries tend to have inadequate surveillance resources, and farmers often have little incentive to report outbreaks because they will not receive any compensation for culled livestock. Countries sometimes also fail to look for, or report, outbreaks so that they can claim they are free of infection and so avoid trade problems.

Flu experts say that the dire state of surveillance could be rapidly turned around by, for example, creating a network of sentinel sites, focusing on the countries and regions most at risk, that would collect isolates and sequence them in real time. Such a network would probably even cost less than the fragmented and uncoordinated surveillance efforts in place today, says Jeremy Farrar, director of the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam (see page 534).

The problem is that no global body has overall responsibility for flu surveillance. The World Health Organization (WHO) runs a global network of labs for human flu surveillance and selects human strains to be included in vaccines for seasonal flu. Monitoring animals falls to the FAO, which tends to focus on food security, and the OIE, which looks mostly at animal health and trade.

What is needed is international leadership, says Farrar. "If, say, the WHO and the FAO were to construct an advisory framework, surveillance could probably be done much more systematically and efficiently."

■ SEE EDITORIAL P. 509

1. Butler, D. & Ruttimann, J. *Nature* **441**, 137–139 (2006).
2. Check, E. *Nature* **442**, 348–350 (2006).
3. Butler, D. *Nature* <http://dx.doi.org/10.1038/news050801-1> (2005).
4. Butler, D. *Nature* <http://dx.doi.org/10.1038/news060206-7> (2006).
5. *Nature* **440**, 255–256 (2006).



Italy's National Galileo Telescope in the Canary Islands will host the HARPS-North planet finder.

SPACE

North set for mass analysis of planets

Spectrograph will review results from Kepler telescope.

BY ERIC HAND

How many extrasolar planets has NASA's Kepler mission discovered? That depends on how you count. Last month, the mission team published a catalogue that lists a staggering 2,321 candidate planets, amassed since May 2009 as the space-based telescope watches stars for the shadow of planets passing over their faces. Yet only 69 of them are considered confirmed planets. Astronomers have fretted over the growing backlog, but help is on its way.

For a Kepler planet to ascend from candidate to confirmed, a second method has to vouch for it: for example, a ground-based spectrograph must report signs that the planet's gravity is tugging its star back and forth. Yet Kepler looks north, whereas the instrument most sensitive to stellar wobbles, the European Southern Observatory's High Accuracy Radial Velocity Planet Searcher (HARPS), is located at the La Silla Observatory in Chile and can only observe the southern sky. On 1 April, however, the Northern Hemisphere will get a near-clone of HARPS when HARPS-North achieves first light at the Italian 3.6-metre National Galileo Telescope (TNG) on La Palma in the Canary Islands.

The instrument has been a long time coming. Conceived in 2005, the project was originally led by Harvard University in Cambridge, Massachusetts. But in 2010, after Harvard's endowment fell during the financial crisis, the University of Geneva in Switzerland took charge. Financial problems forced the group to switch from the 4.2-metre William Herschel Telescope, also on La Palma, to the TNG, which will give the HARPS-North team 80 nights of dedicated time per year for five years.

That should help to alleviate the bottleneck for Kepler candidates. Many astronomers, however, are looking to HARPS-North less for confirmation of the candidate planets than for insight into their properties. The false-positive rate for Kepler, after all, has already been shown to be less than 10% (T. D. Morton and J. A. Johnson *Astrophys. J.* (in the press) Preprint at <http://arxiv.org/abs/1101.5630>; 2011). "It has become acceptable to do a statistical analysis and say, 'They are planets,'" says Joshua Winn, an astronomer at the Massachusetts Institute of Technology in Cambridge.

What Winn and

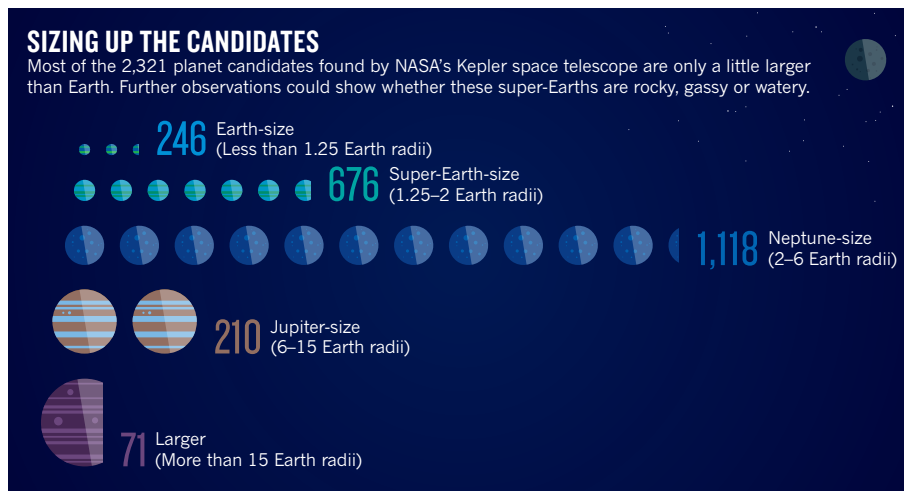
➔ **NATURE.COM**
For more on
Kepler's search for
exoplanets, visit:
go.nature.com/ptgbbt

INAF

others want to know is: what kind of planets? The stellar dimming that Kepler detects is a function of the planet's size. The stellar wobble seen by instruments such as HARPS-North, by contrast, reveals a planet's mass. With knowledge of both mass and size, astronomers can compute the planet's density, which is the key to understanding the nature of the super-Earths — planets a few times the size of Earth (see 'Sizing up the candidates') — that are emerging from Kepler's discoveries. "You really want to know if it's a large Earth or a mini-Neptune or something different altogether, like a ball of water," says David Charbonneau, a Harvard astronomer and a collaborator on the HARPS-North project.

Until now, Kepler has relied for follow-up observations on an instrument on one of the twin 10-metre telescopes at the Keck Observatory atop Mauna Kea in Hawaii, which can detect a stellar wobble of about 150 centimetres per second — barely good enough to confirm super-Earths. HARPS achieves twice that sensitivity, and Francesco Pepe, principal investigator of HARPS-North at the University of Geneva, hopes that his project will do the same.

Kepler's ultimate goal is to detect Earth-sized planets in Earth-like orbits, and for now, such true Earth analogues remain out of reach of ground-based detectors. (The Sun's wobble owing to Earth is about 9 centimetres



per second.) To achieve that sort of precision, astronomers want to fit ground-based spectrographs with high-frequency lasers that emit extremely short pulses of light at specific wavelengths spaced equally on the spectrum. These 'laser frequency combs' provide a way to calibrate a spectrograph and synch it to an atomic clock, which eliminates the error resulting from long-term drift of the spectrograph and could offer precision as good as one centimetre per second.

Ronald Walsworth, a physicist at Harvard,

plans to take his group's laser comb to HARPS-North later this year. A rival group at the European Southern Observatory is already testing a comb at HARPS.

But these efforts may be for nought. The roiling surfaces of stars jitter, creating a confounding noise in the measurement. It may be that only a small fraction of stars are calm enough for an Earth analogue to be discernible, says Walsworth. "Our future is in the stars," he says. "We'll do the science that the stars enable us to do." ■

Chemistry's web of data expands

Patent information to be made publicly accessible amid worries about data quality.

BY RICHARD VAN NOORDEN

Other areas of science feast on free online data, but chemistry has been late to the party. Now it is catching up.

In the latest effort to provide free access to chemical information, the London-based company SureChem (owned by Digital Science, a sister company to Nature Publishing Group) said this week that it would release data on 10 million molecules patented by the pharmaceutical industry since 1976. Harvested automatically from some 20 million patents, the data could lower barriers to drug discovery by academic researchers.

The announcement, made on 26 March at the spring meeting of the American Chemical Society (ACS) in San Diego, California, follows a similar move by computing giant IBM last December. IBM deposited computer-harvested data on about 2.4 million small molecules into PubChem, the world's largest free chemistry repository, which is run by the US National Library of Medicine in Bethesda, Maryland.

Both data releases serve in part to promote the companies' subscription services for patent and structure analysis. But Michael Walters, a chemist working in academic drug discovery at the University of Minnesota in Minneapolis, thinks that the initiatives could mark "a sea change in the way in which patent data are accessed and analysed". The data should make it easier for chemists to see which bioactive molecules have drawn the attention of the drug industry — and to explore new drug targets by designing compounds that are not named in patents.

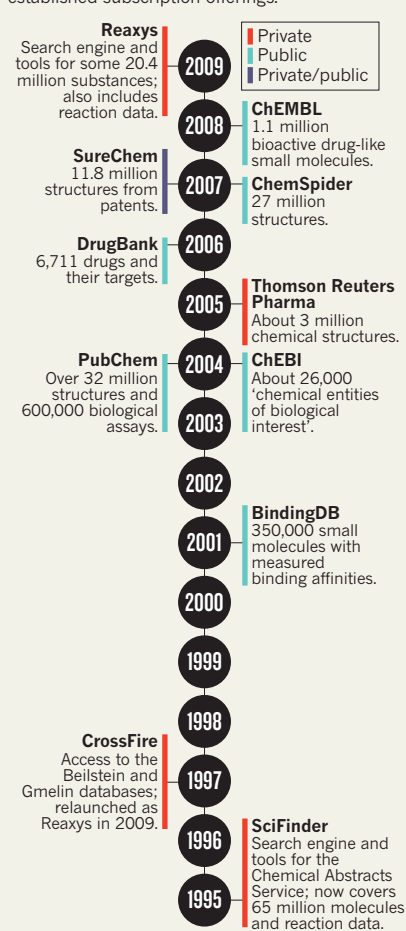
CHEMISTRY ON THE INTERNET

Academic drug discovery will get another boost in September, when a consortium of eight pharmaceutical firms, three biotechnology companies and a number of leading informaticians releases its own free, online drug-discovery platform, the Open Pharmacological Concepts Triple Store (OpenPHACTS). Supported in part by a €10-million (US\$13-million) grant from the European Union's Innovative Medicines Initiative, the website will link data on small molecules and their biological effects, to provide a library of compounds that anyone can download and explore.

Unlike biologists, who are swamped by free databases on genes and proteins, chemists have always expected to pay for their data. Until a few years ago, the market in chemical information was monopolized by the ACS Chemical

CHEMISTRY BREAKS FREE

A web of free, online chemical databases has sprung up over the past decade, alongside established subscription offerings.



to PubChem, has roughly 100,000 unique users per day. In 2007, another free repository, ChemSpider, was created by chemist Antony Williams; in 2009, it was purchased by the UK Royal Society of Chemistry in London and it now holds 27 million structures.

These two databases are now the Internet's main chemistry hubs, linking out to other sources of free online information, such as ChEMBL, a database of about 1 million bioactive drug-like small molecules hosted by the European Bioinformatics Institute in Hinxton, UK. The result is a web of interconnected free data, contrasting with high-quality but closed-off subscription databases.

QUALITY CONTROL

But as biologists already know, free online data can be poorly curated — and chemical data is no exception. In a project presented at the ACS meeting in San Diego, Williams and his colleagues showed how five large online databases disagreed on the structures of 150 top-selling drugs: the best got 99% of structures correct, whereas the worst managed only 76%. In fact, notes Williams, Wikipedia proved the most reliable source of structural information in that experiment — mostly because of an effort to clean up the site's 13,000 pages about chemicals.

Williams says that more chemists need to concentrate on data standards and start actively correcting information online. Christopher Southan, a chemical-information consultant in Gothenburg, Sweden, who previously worked for drug giant AstraZeneca, agrees: "The danger is that now people are connecting all this online chemical and biological information together, and there's so much noise and imprecision that they're building a house of cards."

For now, cheminformatics pioneers are excited by the potential of free online information, and are keen to raise awareness of the possibilities. "The average medicinal chemist was weaned on SciFinder," says Southan. "I can't see them rushing into online data — but slowly but surely, anyone working in academic drug discovery will start to use it." ■

CORRECTION

The graphic 'Frequent fliers' in the News story 'Activists ground primate flights' (*Nature* **483**, 381–382; 2012) should have listed the American Anti-Vivisection Society instead of PETA in its source list.



DIRT POOR

*The key to tackling hunger in Africa is enriching its soil.
The big debate is about how to do it.*

BY NATASHA GILBERT

Eneless Beyadi appears through a forest of maize clutching an armful of vegetables and flashing a broad smile.

Beyadi cultivates about half a hectare of plots in the village of Nankhunda, high on the Zomba plateau in southern Malawi. She gets up at 4 a.m. every day to tend her gardens, as she lovingly calls them, before heading off to teach at a school. In the afternoon, she returns to the gardens, which help to feed her family of six. As testimony to her efforts, the maize (corn) on Beyadi's land stands tall even in the lashing rain, whereas the stunted, yellowed stalks on a neighbour's plot bow low.

The strength of Beyadi's crop is down to more than her green fingers, though. It is also due to what she feeds the soil. Beyadi borrowed money from a European friend to purchase two 50-kilogram bags of chemical fertilizer for this growing season. Because a bag can cost up to 4,000 Malawian kwachas (US\$24), it is beyond the reach of many Malawians, including Beyadi's neighbour, Catharine Changuya, an unmarried mother of four.

Fertilizers make such a profound difference

here because the rusty red soil, as in many parts of Africa, is deficient in organic matter and in key nutrients such as nitrogen and phosphorus. By farming intensively without replenishing soil nutrients, farmers across sub-Saharan Africa have lost an average of 22 kilograms of nitrogen, 2.5 kilograms of phosphorus, and 15 kilograms of potassium per hectare annually over the past 30 years — the yearly equivalent of US\$4 billions' worth of fertilizer. As a result, yields are meagre.

Agricultural experts worry that Africa's soil problems are heading towards a crisis. "The future picture is dire," says Dennis Garrity, chief executive of the World Agroforestry Centre (ICRAF), headquartered in Nairobi. "Producing more food for a growing population in the coming decades, while at the same time combating poverty and hunger, is a huge challenge facing African agriculture."

➔ NATURE.COM
For a podcast and
slideshow on this
story see:
go.nature.com/e7i3ua

African governments, international donors and scientists all agree that farmers must revitalize

their soils. But there is passionate debate about how to do it. Many African governments and agricultural scientists argue that large doses of inorganic fertilizers are the most practical solution. But others, such as the Food and Agriculture Organization of the United Nations (FAO) in Rome, are pushing for greener, cheaper solutions, such as no-till farming that conserves soil and 'fertilizer plants' that boost the soil's nitrogen content organically. Researchers report that these latter techniques are beginning to raise yields and improve soil fertility. But farmers are slow to adopt such practices, which require significantly more labour.

Leading scientific and political figures will take the debate to the UN's Earth Summit in Rio de Janeiro, Brazil, in June. But whatever they recommend, the biggest test is what happens when Beyadi and other African farmers try to put into practice the grand plans of scientists, international donors and governments.

"Many people are promoting approaches

Nitrogen-fixing plants could help improve fertility in African soils.

WOODFALL WILD IMAGES/PHOTO SHOT

without understanding the conditions in Africa and the communities and what works for them. They mean well, but they need to appreciate the realities of the smallholder farmer," says Bashir Jama, director of the soil-health programme for the Alliance for a Green Revolution in Africa (AGRA), based in Nairobi.

Sub-Saharan Africa is one of the poorest regions on Earth, in both living standards and soil fertility. The depleted soil has caused average yields of grain crops to stagnate at around 1 tonne per hectare since the 1960s. By contrast, yields now reach 2.5 t ha⁻¹ in south Asia and 4.5 t ha⁻¹ in east Asia, where chemical fertilizers have been widely adopted since the green revolution (see 'Uneven Landscape'). Fertilizer use across Africa has remained at around 9 kg ha⁻¹ of cultivated land over the past 40 years, whereas Asia uses 96 kg ha⁻¹ of inorganic fertilizer.

Cost is one of the biggest problems. Because of transport expenses, farmers in inland Africa pay more than twice as much for fertilizer as farmers in Europe. And supply is often unreliable because of poor distribution systems.

The World Bank and other major international donors helped to fund fertilizer use in sub-Saharan Africa in the 1970s and 80s, but they came to see such subsidies as a drag on private-sector development and cut them off, pushing African nations to cease offering them as well. However, when Malawi faced a major food crisis in 2005, President Bingu wa Mutharika, who was facing re-election, reintroduced subsidies for fertilizers and improved seeds.

Over the next few years, that policy reaped strong agricultural gains, which came to be known as the Malawi miracle. As fertilizer use in the country almost doubled between 2005 and 2009, maize yields surged from around 1 t ha⁻¹ in 2005 to just under 3 t ha⁻¹ in 2009–10, according to government figures. The agricultural subsidy programme cost the Malawian government \$461.4 million over five years, and comprised 13.5% of the national budget in 2009, the most recent year for which figures are available.

Obtaining accurate data on yields is difficult, says Andrew Dorward, an agro-economist at the School of Oriental and African Studies in London, who has analysed the Malawi subsidy scheme, but he is convinced that the subsidy is having a positive effect. And he is not alone. "Before the subsidy the country was a patchwork of yellow maize. Today it's all green," says Stephen Carr, a retired consultant on African agriculture based in Zomba, who has worked there since the end of the 1980s.

The rise in crop yields helped to convince the World Bank to soften its stance in 2007, when it announced that subsidies "may be justifiable on a temporary basis to stimulate increased fertilizer use in the short term".

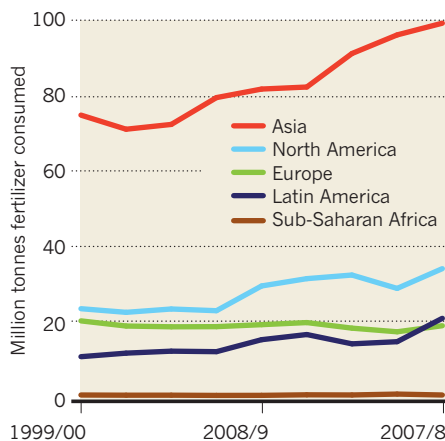
Other African nations have watched Malawi's experiment with interest, and some — including Rwanda, Zambia and Mali — have

ramped up their own schemes. Rwanda began subsidizing fertilizer transport costs in 2006. Fertilizer imports to the country nearly tripled between 2005 and 2007, and wheat and maize yields have increased by 16% and 73%, respectively, over the past five years. The Rwandan system differs from Malawi's by encouraging the development of a fertilizer distribution industry, which is more attractive to donors and is supported by the World Bank.

An international collaboration of researchers is hoping to improve the use of fertilizers by developing digital soil maps covering 42 African countries south of the Sahara. Started in 2009 with an \$18-million grant from AGRA and the Bill & Melinda Gates Foundation to the Tropical Soil Biology and Fertility Institute in Nairobi, the maps will provide up-to-date information on soil properties, derived from satellite measurements and sampling at 60 sites across Africa. Keith Shepherd, a soil scientist at ICRAC who has worked on the maps, says that the analysis will inform agronomists and agricultural extension services about soil health and what nutrients are lacking. "Until now there was no unbiased sampling at this scale so there was no reliable data on acute problems," he says.

UNEVEN LANDSCAPE

Unlike Latin America and Asia, sub-Saharan Africa did not increase its use of fertilizer in the 2000s.



AGRA is also helping to bring some of Africa's better-quality phosphate deposits into production, which will provide sub-Saharan countries with a cheaper source of locally produced phosphate fertilizer.

Even so, fertilizer use in Africa is at the mercy of precarious politics. Although Rwanda's fertilizer programme is growing, Malawi's has started to fall apart as the country's economy has collapsed and its international relations have deteriorated. Many of Malawi's biggest donors, including the UK government's Department for International Development, suspended budgetary support to the nation last year because of concerns about governance and the Malawian government's refusal

to devalue its currency as recommended by the International Monetary Fund.

Although the United Kingdom reinstated some funding to help transport fertilizer, many Malawians couldn't purchase it this year. Changuya walked for an hour and a half to the depot in town, only to find that all the subsidized fertilizer was gone and she would not have been able to afford it anyway.

GREENER SOLUTIONS

With fertilizer-subsidy schemes in trouble, many researchers and donors are supporting more-sustainable methods of boosting yields. They argue that long-term fertilizer use is not only too expensive but also degrades the environment, in particular by releasing the greenhouse gas nitrous oxide.

"The conventional approach has been to focus on improved seeds and chemical fertilizers, but I think there is plenty of evidence to show that there are alternatives that demand significant attention," says Garrity.

One green solution gaining attention is to plant high-protein nitrogen-fixing legumes such as pigeon pea, peanuts and soya beans. With help from bacteria in their roots, legumes capture nitrogen from the atmosphere and convert it into compounds that can be used by plants. They can add up to 300 kg N ha⁻¹ to the soil in a season. Farmers can plant legumes next to grain crops or they can be alternated by season. Studies in experimental plots in Malawi showed that the legumes increased maize yields by 116%.

Although that strategy looks great on paper, poor farmers cannot generally give up much of their limited land to grow legumes, which require extra labour. So, many of the leguminous plants are underperforming in the field. On the average smallholding in Africa, they often fix less than 8 kg N ha⁻¹ per year.

Ken Giller, an agronomist at Wageningen University in the Netherlands, is hoping to tackle these problems through a four-year research programme called N2 Africa, which he helped to start with \$22 million from the Bill & Melinda Gates Foundation and the Howard G. Buffet Foundation. The project aims to breed edible legumes with increased yields and better nitrogen-fixing ability, and to help spread legume crops across Malawi and seven other African nations.

The FAO is promoting other green ways of raising yields, in particular an approach called conservation agriculture, which involves covering fields with mulch and not tilling the soil. The FAO says that this improves soil fertility while reducing erosion and labour. In June 2011, the FAO launched a programme called Save and Grow, currently funded at around \$7 million per year, which over the next 15 years aims to promote research, training and resources for this type of agriculture. But critics argue that conservation agriculture can actually decrease yields and that few farmers

SOURCE: IFDC



Farmer Eneless Beyadi uses inorganic fertilizers and sustainable growing techniques to replenish the nutrients in her plots in southern Malawi.

in sub-Saharan Africa are willing to use these techniques. The promotion of conservation agriculture “is wholly misplaced”, says Giller.

In experimental fields outside Lilongwe in Malawi, researchers are studying another type of green approach as part of an ICRAF project run by Gudeta Sileshi, an agricultural scientist at the Chitedze Agricultural Research Station. At the far end of one field, mature *Faidherbia albida* trees stand some 4 metres high, their leafless branches feathery against the sky. Smaller, younger trees dot the fields. They are all ‘fertilizer trees’, which fix nitrogen and improve the nitrogen content of the soil. Past agricultural-improvement schemes have tried unsuccessfully to sell the idea of these trees to African farmers, but there is renewed interest from researchers and governments, because tree planting is seen as one way of sequestering carbon and combating global warming.

Sileshi is testing several types of trees, but he is betting that the top performer will be *F. albida*, which is indigenous to Africa. Besides fixing nitrogen, it has deep roots that draw nutrients from far below the surface and store them in the tree’s spiky leaves. When the leaves fall, nitrogen is returned to the top layer of the soil for use by crops planted beneath the tree’s canopy. *Faidherbia albida* sheds its leaves early in the rainy season when crops are beginning to grow, so it doesn’t compete with them for light, nutrients or water; its summer canopy of leaves reduces transpiration from crops underneath it and evaporation from the soil. In a trial in Zambia, maize yields under *F. albida* reached 4 t ha⁻¹ compared with 1.3 t ha⁻¹ outside the canopy.

In the Chimalanga 2 village near the shores

of Lake Malawi, a healthy crop of maize grows beneath the naked branches of *F. albida*. “If everyone planted trees, we would reduce the problem and return the soil to how it used to be,” says Beather Kandaya, a farmer in the village. Kandaya was trained in fertilizer-tree husbandry by Malawi’s agricultural extension service and she passes the knowledge on to her neighbours. But the practice is not catching on quickly. In Malawi, just over 1% of farmers grow the trees.

This could soon change. Malawi, Niger, Kenya and Rwanda are among the African countries promoting the use of trees within

“The fertilizer subsidy is a matter of life and death.”

farming. And international donors are starting to support tree programmes.

Beyadi’s experience, though, shows how much work there is still to do. She and her neighbours care for a small experimental nursery filled with some 200 spindly seedlings of leguminous plants and trees. Of the 240 *F. albida* seeds that the group planted last year, only one seedling has survived. It is a common problem because the roots of *F. albida* are very fragile and easily damaged. And it will be 6–10 years before an *F. albida* sapling is ready to fertilize crops. At the Chitedze research station, ICRAF and the agricultural non-governmental organization Total Land Care, based in Lilongwe, are trying to find ways of cultivating

F. albida more efficiently.

Proponents of fertilizer trees say that they deserve a chance and that they have received far less funding and promotion than other sustainable farming techniques and inorganic fertilizers. “The simplest, cheapest techniques are the ones that are ignored as they are not of interest to the private sector,” says Garrity. “It’s time we no longer ignored them.”

But many agricultural experts and farmers conclude that green approaches are not enough — that Africa can’t solve its soil problems without chemical fertilizers. “I continue to believe leguminous trees and plants have got to play a part in maintaining soil fertility and food security, but they can’t replace inorganic fertilizer,” Carr says. “The fertilizer subsidy is a matter of life or death.”

Giller argues that to win over farmers, new techniques for increasing crop yield must bring extra benefits, which is why he is focusing his research on edible legumes. The acid test is whether farmers continue to use the new approaches after aid support ends.

Beyadi acknowledges that many of her fellow farmers will drop the new green techniques when aid goes. And although she says that she will continue to care for the fertilizer trees, she doesn’t regard them as her best hope for the future. Like her neighbours, she sees inorganic fertilizers as the key to growing more food. Surveying her lush green gardens, Beyadi wonders whether she will be able to buy enough fertilizer next year to ensure an equally bountiful harvest. ■ [SEE EDITORIAL P. 510.](#)

Natasha Gilbert is a reporter for Nature in London.



THE BIOLOGICAL HIGGS

Biologists ponder what fundamental discoveries might match the excitement of the Higgs boson.

BY HEIDI LEDFORD

Biologists may have little cause to envy physicists — they generally enjoy more generous funding, more commercial interest and more popular support. But they could have been forgiven a moment of physics envy last December when, after a week of build-up and speculation, researchers at the Large Hadron Collider (LHC) near Geneva in Switzerland addressed a tense, standing-room-only auditorium.

Scientists there had caught the strongest hints yet of the Higgs boson: what some have called the ‘God particle’ and the final missing piece of the standard model that explains the behaviour of subatomic particles. The discovery, if confirmed, will mark the culmination of a hunt that has taken years and cost billions of dollars, and will shape the field for years to come. The research community was abuzz. “There were lots of rumours flying around about how significant the signal was,” says Lisa Randall, a theoretical particle physicist at Harvard University in

Cambridge, Massachusetts, who got up at 4 a.m. to talk to the press before watching the webcast of the presentation at the LHC. “It’s been quite exciting.”

All this led *Nature* to wonder: what fundamental discoveries in biology might inspire the same thrill? We put the question to experts in various fields. Biology is no stranger to large, international collaborations with lofty goals, they pointed out — the race to sequence the human genome around the turn of the century had scientists riveted. But most biological quests lack the mathematical precision, focus and binary satisfaction of a yes-or-no answer that characterize the pursuit of the Higgs. “Most of what is important is messy, and not given to a moment when you plant a flag and crack the champagne,” says Steven Hyman, a neuroscientist at the Broad Institute in Cambridge, Massachusetts.

Nevertheless, our informal survey shows that the field has no shortage of fundamental questions that could fill an anticipatory auditorium. These questions concern where and how life started — and why it ends.

ILLUSTRATIONS BY B. MONROE

IS THERE LIFE ELSEWHERE?

In 1964, palaeontologist George Gaylord Simpson wrote a stinging dismissal of exobiology, the search for life on other planets. “This ‘science’ has yet to demonstrate that its subject matter exists!” he wrote¹. The searing critique caused many researchers in the nascent field to shy away from exobiology.

But it was unfair, says planetary scientist Christopher Chyba of Princeton University in New Jersey. Chyba has for years been comparing the search for life on other planets to the search for the Higgs: another quest whose subject has never been proved to exist. “Why should we suddenly become giggly when it is biology at stake, rather than physics?” Chyba wrote in a 2005 rebuttal to Simpson’s attack².

The search for extraterrestrial life can be described as one way to test “a standard model of biology”, says astrobiologist Chris McKay of the NASA Ames Research Center in Moffett Field, California. “It’s the model of DNA and amino acids and proteins and a genetic code,” he says. “It’s the common features of all biology, and the framework through which everything we know about life is based.” If life fundamentally different from this standard model — perhaps relying on a wildly different biochemistry — were found on another planet, it would show that there is more than one way to produce a living system, he adds.

Others say they don’t need evidence of such a ‘second genesis’ to get a Higgs-like thrill from the prospect of life on other planets. “If we found our same biology, but on Mars, that would be pretty exciting,” says biochemist Gerald Joyce of the Scripps Research Institute in La Jolla, California. “Then the question would be: where did it come from first?”

But whereas the Higgs-hunters in Geneva have a good idea of what to look for, astrobiologists seeking alternative forms of life face a bigger logistical challenge: figuring out what clues are most revealing. The chemical signatures of compounds that are commonly associated with life, such as methane or liquid water, could identify planets to focus on. But atmospheric signatures of life are unlikely to be convincing, says Chyba.

Within the Solar System, McKay puts his money on three habitats as most likely to harbour life: Enceladus, an icy moon orbiting Saturn that, according to NASA’s Cassini spacecraft, probably has liquid water and is spewing organic material from cracks in its surface³; Mars, but “old Mars, not Mars today”; and Jupiter’s moon, Europa, whose icy surface masks tantalizing seas of water. The Mars Science Laboratory, scheduled to land on the red planet in August, will include a simple mass spectrometer and a laser spectrometer, enabling it to detect methane, and could reveal preliminary signs of life. But the mission is not designed to yield definitive evidence.

Another way to hunt for life is to look for organic molecules that are too complex to have arisen by simple chemical synthesis, unaided by enzymes. “Let’s say you came to Earth and scooped up matter,” says McKay. “You’d find all of this chlorophyll and DNA: big, huge, complex molecules that were clearly there in high abundance and distinctly different from what you’d expect from a chemical mix.” Finding this would require sophisticated equipment that had been baked and scrubbed free of earthly contaminants and, at present, there are no concrete plans to include such equipment on NASA’s proposed trips to Mars or Europa. “My sense is that people are just trying to avoid it

as long as possible,” Chyba says. “Money is extremely tight, but at some point we’ll just have to bite the bullet.”

Searching rocks on other planets for fossils is another popular proposition, says Jeffrey Bada, a planetary geochemist at the Scripps Institution of Oceanography in La Jolla. “That’s easy enough,” he says. “But if you don’t find them, does that tell you that life never existed there?” McKay argues that fossil evidence or living proof of life may be required to convince a field. “Ultimately, you’ll have to have a body,” he says. “It doesn’t have to be alive, but you’ll have to have a body.”

IS THERE FOREIGN LIFE ON EARTH?

Alien life — and a Higgs moment — might also be lurking close to home. Some have postulated the existence of a ‘shadow biosphere’ on Earth, teeming with life that has gone undiscovered because scientists simply don’t know where to look. It could contain life that relies on a fundamentally different biochemistry, using different forms of amino acids or even entirely novel ways of storing, replicating and executing inherited information that do not rely on DNA or proteins.

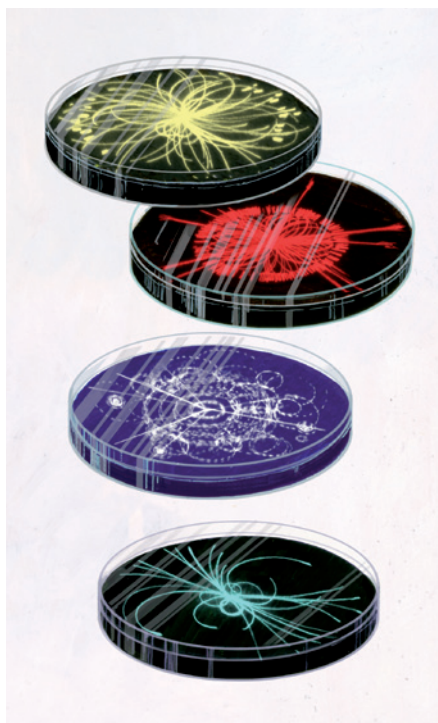
The idea is not as far-fetched as it might sound, says Steven Benner, a chemist at the Foundation for Applied Molecular Evolution in Gainesville, Florida. Researchers have found shadow biospheres before. The invention of the microscope revealed whole new worlds, says Benner; and the discovery of a new realm of microorganisms, the archaea, opened a window on another. “The question is: is it going to happen again?”

The trick is deciding what to look for and how to detect it. The usual way that researchers search for new organisms — by sequencing DNA or RNA — will not pick up life that does not depend on them.

Some scientists have speculated that desert varnish, a peculiar dark-coloured coating of unknown origin found on many desert rocks, could be a product of a shadow biosphere. Benner suggests looking in nooks and crannies that cannot support conventional life, such as areas with extremely high temperatures, radiation levels or harsh chemical environments.

Felisa Wolfe-Simon, now at the Lawrence Berkeley National Laboratory in Berkeley, California, and her colleagues took this approach when they searched for life in the arsenic-rich environment of California’s Mono Lake. In late 2010, they reported the discovery of a life form that can use arsenic in place of phosphorus in its DNA and proteins — a seemingly remarkable departure from conventional life⁴. But at least one attempt to reproduce the result has failed.

Another approach is to search on the basis of size. If cells were liberated from their reliance on bulky ribosomes and proteins, they could be much smaller, says Benner, perhaps tucked away in rocks with pores only nanometres across. That is the rationale behind a project that John Atkins, a molecular geneticist at the University of Utah in Salt Lake City, is pursuing with Richard Herrington of the Natural History Museum in London. They plan to sequence the contents of rocks of different ages and origins with pores less than 100 nanometres in diameter. By screening for nucleic-acid sequences that lack the code for protein-making ribosomes, they hope to find a protein-free life form that has its roots in RNA, as known life probably does, but that arose independently. “The RNA world is thought to have originated, in geological terms, relatively quickly,” Atkins says. “So why couldn’t it have arisen again multiple times?”



HOW DID LIFE START ...?

Even if alternative forms of life elude scientists, a fuller picture of how familiar life originated on Earth would surely create ripples in biology.

Joyce says that there will come a point at which researchers learn how to synthesize an evolving, replicating system from scratch. Getting there won't have the "monolithic, big-science march across the goal line" that has characterized the search for the Higgs, he cautions. But it will answer a key biological question: what does it take to create life from a primordial soup? And that could provide insight into how life on Earth began. "We'll never know for sure, but at least you can test plausible hypotheses," says James Collins, a synthetic biologist at Boston University in Massachusetts.

Several labs have already made headway. Joyce and his collaborators have pioneered work on the RNA-world concept, in which RNA molecules, capable of encoding information and catalysing chemical reactions, replicated and evolved faster than they degraded. RNA is notoriously unstable, and the idea is that, over time, this system gave way to DNA, a sturdier system for storing information, and proteins, a more versatile mode of catalysing reactions. "The transition to DNA and protein created the potential to evolve into more complex things," Bada says.

In 2009, a paper from Joyce's lab reported the development of a system of RNA molecules that undergo self-sustaining Darwinian evolution⁵. But enzymes and a human hand were needed to create the RNA sequences to start off the reaction, Joyce says, and so far his lab has not found conditions that would allow the system to form spontaneously. "We're still a bit challenged," he says. "But the system is running more and more efficiently all the time."

Jack Szostak and his colleagues at Harvard Medical School in Boston have taken a different approach, enclosing RNA molecules in fatty-acid vesicles as an early step towards the creation of a primitive cell. The vesicles grow and divide spontaneously, but the genetic material does not replicate without the aid of an enzyme⁶.

Some believe that RNA may have had a precursor. Ramanarayanan Krishnamurthy at the Scripps Research Institute, is testing novel polymers of organic chemicals that could have formed in the primordial goo, in search of those that could replicate and evolve. "RNA was not the first living entity," says Bada. "It's too complex. Something preceded RNA, and that's where the interest is right now."

... AND CAN WE DELAY ITS END?

In a 1993 review⁷, Linda Partridge and Nicholas Barton, both then researchers on ageing at the University of Edinburgh, UK, delivered "a baleful message" to the field of gerontology. The complexity of the biological networks that influence ageing, they wrote, means "it is most unlikely that engineering of a few genes or intervention in a handful of physiological pathways will prevent the process from occurring".

Things have changed. "I could tolerate that debate 20 years ago," says Richard Miller, who studies ageing at the University of Michigan in Ann Arbor. "But now it's just wrong."

Some eight months after the publication of Partridge and Barton's review, Cynthia Kenyon and her colleagues at the University of California, San Francisco, reported that mutations in a single gene allowed the nematode *Caenorhabditis elegans* to live more than twice as long as usual⁸. Three years later, a group led by Andrzej Bartke, who studies ageing at Southern Illinois University in Springfield, reported that mice bearing a single mutation that causes hormonal deficiencies live up to 68% longer than mice without the mutation⁹.

Both papers, and a slew of work since, have suggested that it might be possible to

significantly slow human ageing and its associated diseases. Such an intervention could have a tremendous impact on society, adding years of health and economic productivity, but creating new strains on a society having to support many more older people. And scientifically, the ability to slow ageing would address Higgs-like fundamental questions about human life: why do we age; what pathways control it; and what are the consequences if they are switched off?

There are signs that such interventions may exist. In 2010, Miller and his colleagues showed that feeding mice a drug called rapamycin lengthened their average lifespan by 10% for males and 18% for females¹⁰. And slashing calorie intake by 25–40% can extend lifespan in mice and other mammals. But there is no proof that these approaches would work in humans and, even if they did, neither is likely to catch on: rapamycin can suppress the immune system, and few people can tolerate brutal dietary restriction.

One major challenge for the field is to prove that a putative life-extending agent actually works — something that in humans would take 60 years or more. Jay Olshansky, who studies ageing at the University of Illinois in Chicago, says the field should set a concrete goal:

a seven-year delay in the onset and progression of age-related disease. "If you look at the risk of most of the things that go wrong with us as we grow older, age-related risk doubles roughly every seven years," he says. "If you eliminate one doubling, you reduce the risk of everything by half. It would be monumental."

Miller has a different goal. "We will have the answer when we have something that we can put in dog food that extends the average dog's lifespan by 15 to 20%," he says. Dogs offer an ideal intermediate between mice and humans, says Miller: they are considered a long-lived species and live side-by-side with humans.

But Partridge and Barton's observations about the complexity of ageing still hold true. Most researchers acknowledge that they are only beginning to understand the molecular networks that regulate ageing and its associated diseases. "I don't believe there's one cause of ageing," says Brian Kennedy, president of the Buck Institute for Research on Aging in Novato, California. "But there are pathways that are designed to modulate many things at one time. I think a lot of the genes and drugs we're studying are tapping into those."

At this point, a life-extending therapy seems a much more distant prospect than does confirmation of the Higgs boson. Last month, researchers announced a bump in data from the Tevatron, the US particle collider at Fermilab in Batavia, Illinois, that is consistent with results from the LHC. It has added to physicists' excitement that they are on the threshold of discovery.

Ageing, however, "is almost the complete inverse of the situation of the Higgs particle", reflects Thomas Kirkwood, a leader in the field at Newcastle University, UK. "Everything that we're learning tells us it's highly unlikely that we'll find a single unitary cause." ■

Heidi Ledford writes for Nature from Cambridge, Massachusetts.

"WHY SHOULD WE SUDDENLY BECOME GIGGLY WHEN IT IS BIOLOGY AT STAKE, RATHER THAN PHYSICS?"

➔ NATURE.COM
For more on the Higgs boson, see the LHC special at: go.nature.com/uzvtfd

1. Simpson, G. G. *Science* **143**, 769–775 (1964).
2. Chyba, C. F. *Science* **308**, 495–496 (2005).
3. Postberg, F., Schmidt, J., Hillier, J., Kempf, S. & Srama, R. *Nature* **474**, 620–622 (2011).
4. Wolfe-Simon, F. et al. *Science* **332**, 1163–1166 (2011).
5. Lincoln, T. A. & Joyce, G. F. *Science* **323**, 1229–1232 (2009).
6. Mansy, S. S. & Szostak, J. W. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 47–54 (2009).
7. Partridge, L. & Barton, N. H. *Nature* **362**, 305–311 (1993).
8. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. *Nature* **366**, 461–464 (1993).
9. Brown-Borg, H. M., Borg, K. E., Meliska, C. J. & Bartke, A. *Nature* **384**, 33 (1996).
10. Miller, R. A. et al. *J. Gerontol. A Biol. Sci. Med. Sci.* **66**, 191–201 (2011).

COMMENT

AVIAN INFLUENZA Shift expertise to track mutations where they emerge **p.534**

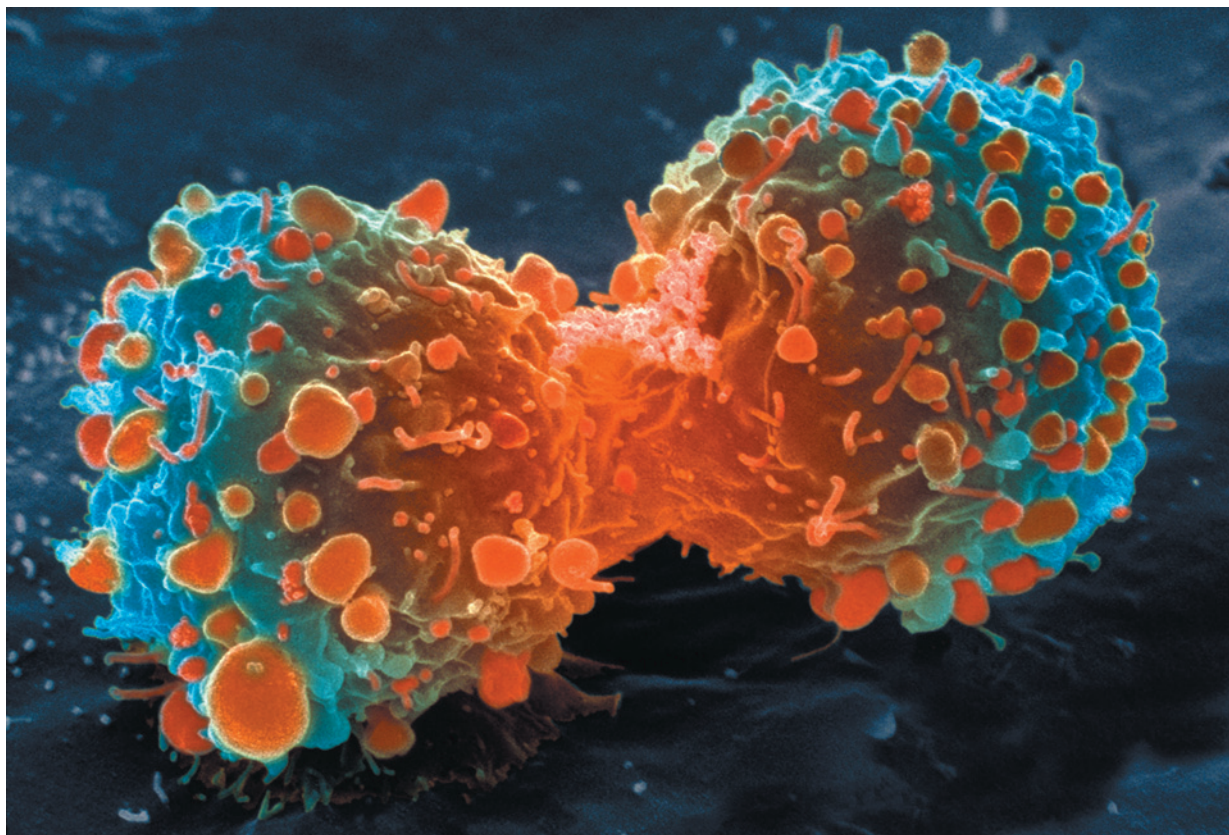
EARTH SYSTEMS Past climates give valuable clues to future warming **p.537**

HISTORY OF SCIENCE Descartes' lost letter tracked using Google **p.540**

OBITUARY Wylie Vale and an elusive stress hormone **p.542**



S. GSCHMEISSNER/SPL



Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low¹. Sadly, clinical

trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and

investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer-cell-line and mouse models² make it difficult for even ►

► the best scientists working in optimal conditions to make a discovery that will ultimately have an impact in the clinic. Issues related to clinical-trial design — such as uncontrolled phase II studies, a reliance on standard criteria for evaluating tumour response and the challenges of selecting patients prospectively — also play a significant part in the dismal success rate³.

Unquestionably, a significant contributor to failure in oncology trials is the quality of published preclinical data. Drug development relies heavily on the literature, especially with regards to new targets and biology. Moreover, clinical endpoints in cancer are defined mainly in terms of patient survival, rather than by the intermediate endpoints seen in other disciplines (for example, cholesterol levels for statins). Thus, it takes many years before the clinical applicability of initial preclinical observations is known. The results of preclinical studies must therefore be very robust to withstand the rigours and challenges of clinical trials, stemming from the heterogeneity of both tumours and patients.

CONFIRMING RESEARCH FINDINGS

The scientific community assumes that the claims in a preclinical study can be taken at face value — that although there might be some errors in detail, the main message of the paper can be relied on and the data will, for the most part, stand the test of time. Unfortunately, this is not always the case. Although the issue of irreproducible data has been discussed between scientists for decades, it has recently received greater attention (see go.nature.com/q7i2up) as the costs of drug development have increased along with the number of late-stage clinical-trial failures and the demand for more effective therapies.

Over the past decade, before pursuing a particular line of research, scientists (including C.G.B.) in the haematology and oncology department at the biotechnology firm Amgen in Thousand Oaks, California, tried to confirm published findings related to that work. Fifty-three papers were deemed 'landmark' studies (see 'Reproducibility of research findings'). It was acknowledged from the outset that some of the data might not hold up, because papers were deliberately selected that described something completely new, such as fresh approaches to targeting cancers or alternative clinical uses for existing therapeutics. Nevertheless, scientific findings were confirmed in only 6 (11%) cases. Even knowing the limitations of preclinical research, this was a shocking result.

Of course, the validation attempts may have failed because of technical differences or difficulties, despite efforts to ensure that this was not the case. Additional models were also used in the validation, because

to drive a drug-development programme it is essential that findings are sufficiently robust and applicable beyond the one narrow experimental model that may have been enough for publication. To address these concerns, when findings could not be reproduced, an attempt was made to contact

"The scientific process demands the highest standards of quality, ethics and rigour."

the original authors, discuss the discrepant findings, exchange reagents and repeat experiments under the authors' direction, occasionally even in the laboratory of the original investigator.

These investigators were all competent, well-meaning scientists who truly wanted to make advances in cancer research.

In studies for which findings could be reproduced, authors had paid close attention to controls, reagents, investigator bias and describing the complete data set. For results that could not be reproduced, however, data were not routinely analysed by investigators blinded to the experimental versus control groups. Investigators frequently presented the results of one experiment, such as a single Western-blot analysis. They sometimes said they presented specific experiments that supported their underlying hypothesis, but that were not reflective of the entire data set. There are no guidelines that require all data sets to be reported in a paper; often, original data are removed during the peer review and publication process.

Unfortunately, Amgen's findings are consistent with those of others in industry. A team at Bayer HealthCare in Germany last year reported⁴ that only about 25% of published preclinical studies could be validated to the point at which projects could continue. Notably, published cancer research represented 70% of the studies analysed in that report, some of which might overlap with the 53 papers examined at Amgen.

Some non-reproducible preclinical papers had spawned an entire field, with hundreds of secondary publications that expanded on elements of the original observation, but did not actually seek to confirm or falsify its fundamental basis. More troubling, some of the research has triggered a series of clinical studies — suggesting that many patients had

subjected themselves to a trial of a regimen or agent that probably wouldn't work.

These results, although disturbing, do not mean that the entire system is flawed. There are many examples of outstanding research that has been rapidly and reliably translated into clinical benefit. In 2011, several new cancer drugs were approved, built on robust preclinical data. However, the inability of industry and clinical trials to validate results from the majority of publications on potential therapeutic targets suggests a general, systemic problem. On speaking with many investigators in academia and industry, we found widespread recognition of this issue.

IMPROVING THE PRECLINICAL ENVIRONMENT

How can the robustness of published preclinical cancer research be increased? Clearly there are fundamental problems in both academia and industry in the way such research is conducted and reported. Addressing these systemic issues will require tremendous commitment and a desire to change the prevalent culture. Perhaps the most crucial element for change is to acknowledge that the bar for reproducibility in performing and presenting preclinical studies must be raised.

An enduring challenge in cancer-drug development lies in the erroneous use and misinterpretation of preclinical data from cell lines and animal models. The limitations of preclinical cancer models have been widely reviewed and are largely acknowledged by the field. They include the use of small numbers of poorly characterized tumour cell lines that inadequately recapitulate human disease, an inability to capture the human tumour environment, a poor appreciation of pharmacokinetics and pharmacodynamics, and the use of problematic endpoints and testing strategies. In addition, preclinical testing rarely includes predictive biomarkers that, when advanced to clinical trials, will help to distinguish those patients who are likely to benefit from a drug.

Wide recognition of the limitations in preclinical cancer studies means that business as usual is no longer an option. Cancer researchers must be more rigorous in their approach to preclinical studies. Given the inherent difficulties of mimicking the human micro-environment in preclinical research, reviewers and editors should demand greater thoroughness.

REPRODUCIBILITY OF RESEARCH FINDINGS

Preclinical research generates many secondary publications, even when results cannot be reproduced.

Journal impact factor	Number of articles	Mean number of citations of non-reproduced articles*	Mean number of citations of reproduced articles
>20	21	248 (range 3–800)	231 (range 82–519)
5–19	32	169 (range 6–1,909)	13 (range 3–24)

Results from ten-year retrospective analysis of experiments performed prospectively. The term 'non-reproduced' was assigned on the basis of findings not being sufficiently robust to drive a drug-development programme.

*Source of citations: Google Scholar, May 2011.

RECOMMENDATIONS

Improving the reliability of preclinical cancer studies

We recommend the following steps to change the culture of oncology research and improve the relevance of translational studies:

- There must be more opportunities to present negative data. It should be the expectation that negative preclinical data will be presented at conferences and in publications. Preclinical investigators should be required to report all findings, regardless of the outcome. To facilitate this, funding agencies, reviewers and journal editors must agree that negative data can be just as informative as positive data.
- Journal editors must play an active part in initiating a cultural change. There must be mechanisms to report negative data that are accessible through PubMed or other search engines. There should be links to journal articles in which investigators have reported alternative findings to those in an initial (sometimes considered landmark) publication. One suggestion is to include 'tags' that report whether the key findings of a seminal paper were confirmed.
- There should be transparent opportunities for trainees, technicians and colleagues to discuss and report troubling or unethical behaviours without fearing adverse consequences.

- Greater dialogue should be encouraged between physicians, scientists, patient advocates and patients. Scientists benefit from learning about clinical reality. Physicians need better knowledge of the challenges and limitations of preclinical studies. Both groups benefit from improved understanding of patients' concerns.
- Institutions and committees should give more credit for teaching and mentoring: relying solely on publications in top-tier journals as the benchmark for promotion or grant funding can be misleading, and does not recognize the valuable contributions of great mentors, educators and administrators.
- Funding organizations must recognize and embrace the need for new cancer-research tools and assist in their development, and in providing greater community access to those tools. Examples include support for establishing large cancer cell-line collections with easy investigator access (a simple, universal material-transfer agreement); capabilities for genetic characterization of newly derived tumour cell lines and xenografts; identification of patient selection biomarkers; and generation of more robust, predictive tumour models. **C.G.B. and L.M.E.**

As with clinical studies, preclinical investigators should be blinded to the control and treatment arms, and use only rigorously validated reagents. All experiments should include and show appropriate positive and negative controls. Critical experiments should be repeated, preferably by different investigators in the same lab, and the entire data set must be represented in the final publication. For example, showing data from tumour models in which a drug is inactive, and may not completely fit an original hypothesis, is just as important as showing models in which the hypothesis was confirmed.

Studies should not be published using a single cell line or model, but should include a number of well-characterized cancer cell lines that are representative of the intended patient population. Cancer researchers must commit to making the difficult, time-consuming and costly transition towards new research tools, as well as adopting more robust, predictive tumour models and improved validation strategies. Similarly, efforts to identify patient-selection biomarkers should be mandatory at the outset of drug development.

Ultimately, however, the responsibility

for design, analysis and presentation of data rests with investigators, the laboratory and the host institution. All are accountable for poor experimental design, a lack of robust supportive data or selective data presentation. The scientific process demands the highest standards of quality, ethics and rigour.

BUILDING A STRONGER SYSTEM

What reasons underlie the publication of erroneous, selective or irreproducible data? The academic system and peer-review process tolerates and perhaps even inadvertently encourages such conduct⁵. To obtain funding, a job, promotion or tenure, researchers need a strong publication record, often including a first-authored high-impact publication. Journal editors, reviewers and grant-review committees often look for a scientific finding that is simple, clear and complete — a 'perfect' story. It is therefore tempting for investigators to submit selected data sets for publication, or even to massage data to fit the underlying hypothesis.

But there are no perfect stories in biology. In fact, gaps in stories can provide opportunities for further research — for example, a treatment that may work in only some cell

lines may allow elucidation of markers of sensitivity or resistance. Journals and grant reviewers must allow for the presentation of imperfect stories, and recognize and reward reproducible results, so that scientists feel less pressure to tell an impossibly perfect story to advance their careers.

Although reviewers, editors and grant-committee members share some responsibility for flaws in the system, investigators must be accountable for the data they generate, analyse and submit. We in the field must remain focused on the purpose of cancer research: to improve the lives of patients. Success in our own careers should be a consequence of outstanding research that has an impact on patients.

The lack of rigour that currently exists around generation and analysis of preclinical data is reminiscent of the situation in clinical research about 50 years ago. The changes that have taken place in clinical-trials processes over that time indicate that changes in prevailing attitudes and philosophies can occur (see 'Improving the reliability of pre-clinical cancer studies').

Improving preclinical cancer research to the point at which it is reproducible and translatable to clinical-trial success will be an extraordinarily difficult challenge. However, it is important to remember that patients are at the centre of all these efforts. If we in the field forget this, it is easy to lose our sense of focus, transparency and urgency. Cancer researchers are funded by community taxes and by the hard work and philanthropic donations of advocates. More importantly, patients rely on us to embrace innovation, make advances and deliver new therapies that will improve their lives. Although hundreds of thousands of research papers are published annually, too few clinical successes have been produced given the public investment of significant financial resources. We need a system that will facilitate a transparent discovery process that frequently and consistently leads to significant patient benefit. ■ **SEE EDITORIAL P.509**

C. Glenn Begley is a consultant and former vice-president and global head of Hematology and Oncology Research at Amgen, Thousand Oaks, California 91359, USA. **Lee M. Ellis** is at the University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, USA. e-mail: lellis@mdanderson.org

1. Hutchinson, L. & Kirk, R. *Nature Rev. Clin. Oncol.* **8**, 189–190 (2011).
2. Francia, G. & Kerbel, R. S. *Nature Biotechnol.* **28**, 561–562 (2010).
3. Rubin, E. H. & Gilliland, D. G. *Nature Rev. Clin. Oncol.* <http://dx.doi.org/10.1038/nrclinonc.2012.22> (2012).
4. Prinz, F., Schlange, T. & Asadullah, K. *Nature Rev. Drug Discov.* **10**, 712 (2011).
5. Fanelli, D. *PLoS ONE* **5**, e10271 (2010).



Investing in the monitoring of new infections in Asia would speed public-health and clinical responses.

Shift expertise to where it matters

Tools and training for responding to diseases such as avian flu must relocate to countries where infections are most likely to emerge, says **Jeremy Farrar**.

Every day on my way to the hospital, I pass streets lined with poultry. The birds disappeared a few years ago, but have gradually returned. This would be of little concern if I was in Europe or North America — but I work in Vietnam studying emerging pathogens, including the avian influenza virus (H5N1). There is a patient with H5N1 in the hospital as I write. This is a region where the H5N1 virus has killed millions of birds and several people, where SARS and Nipah virus emerged and where the threat of antibiotic and antimalarial drug resistance is growing. Because of this, we have become acutely aware of the continued danger of infectious diseases, and the inadequacies of our current systems for tracking

them and responding in a timely fashion.

Too often, surveillance is crisis-driven, ad hoc and reactive; it is incorporated into overextended and under-resourced systems. It frequently relies on outside experts, who arrive with little understanding or appreciation of the country, local infrastructure or culture. Inevitably, a lot of time and resources get wasted — a purchased PCR machine ends up collecting dust in an empty lab, commitments to supply consumables are not honoured and there are too few trained people to service and run things after short-term projects end. Much donor funding is wasted on meetings, teleconferences, workshops and flying in consultants.

This disconnect between the people and

places involved continues to create problems. In 2007, Indonesia put on hold the sharing of H5N1 samples with the World Health Organization, out of a concern that they would be used to create vaccines and other therapies that only wealthy countries could afford. No one can condone refusing to share information of such public-health importance, but today, eight years after Vietnam's first human case of H5N1, too few endemic countries have access to the vaccines or intravenous antiviral drugs that continue to be stockpiled in richer parts of the world. Recent research has pinpointed specific mutations in the H5N1 virus that may render it more transmissible in mammals — but scientists like us in endemic areas are still waiting to learn what those mutations are.

ON THE SPOT

With moderate investment, we could be conducting surveillance for H5N1 and other emerging infections much more effectively, and could link that surveillance with immediate action. Surveillance on its own without a public-health need or clinical response is of questionable value, and unlikely to be sustained. I believe that we have to bring some of the huge investment by the developed world in genomics, technology and training to affected countries in Asia and elsewhere.

In this way, surveillance, analysis of samples, and — crucially — the public-health and clinical-research response can be conducted in the same place, making the process faster and more flexible in dealing with rapid developments. It would require a transfer of technology, prolonged exchange of scientists and a sustained commitment to investment and training locally — along with an equitable sharing of the benefits of the research.

The unit in which I work in Vietnam shows that this type of project is possible. Over the past 21 years, we — alongside our sister programme in Thailand and with partnerships across Asia — have helped to train thousands of regional scientists in clinical medicine, epidemiology, microbiology, bioinformatics and other disciplines crucial to monitoring, controlling and understanding infectious diseases and outbreaks. We are small and flexible, which keeps bureaucracy and costs down — we employ only a few hundred staff across several countries, but collaborate with many more. Thanks to funding from the Vietnamese government, the UK Wellcome Trust, the US National Institutes of Health and the Li Ka Shing Foundation in Hong Kong, we have some

“There is now a window of opportunity to build global scientific capacity before another crisis hits.”

of the capacity and flexibility needed to respond immediately to the rapidly changing dynamic of infectious diseases such as H5N1, enterovirus 71 or artemisinin-resistant malaria, and can make results available in real time. Such an approach is impossible when the work requires individuals to fly in and out and to analyse samples in another country.

SMALL BUT POWERFUL

There are other great examples of long-term research partnerships between national and international organizations, but they are all too few. These infrastructures are easier to build than many believe — you need only a small group of committed people, a shared vision and ethos, flexible funding that encourages local decision-making, and a focus on excellence. There can be great power in such small institutions — which may need as little as a few hundred thousand dollars a year to operate — if only we made better use of them (G. T. Keusch and C. A. Medlin *Nature* **422**, 561–562; 2003).

Because our research unit is based in the region where the story is unfolding, we can appreciate the social issues that can stymie even the best scientific endeavour. For instance, small-scale backyard poultry farms (often family farms with mixed chickens, ducks and pigs) remain a crucial livelihood and the main source of protein for many households in rural Asia. Because no adequate compensation schemes have been developed to encourage reporting of sick poultry and livestock, the usual responses are to cull all local poultry and apportion blame. Such activities can ruin small farmers and their families.

There is now a window of opportunity to build global scientific capacity before another crisis — such as a new pandemic — hits. This means collaborating with the people who share a vested interest in using the money efficiently and effectively to prevent outbreaks and address daily public-health and clinical issues in their own countries. After living in Vietnam for more than 16 years and raising my family here, I can understand the feeling of urgency. Everyone I work with who sees chickens each day on their way to work, hears about local outbreaks in the news or treats patients is united in the effort to stay one step ahead of H5N1 and other potentially deadly outbreaks. We must share the available knowledge and the tools to make it possible — an undertaking that will require us to shift the centre of gravity for such research to where the needs are greatest. ■

Jeremy Farrar is at the Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. e-mail: jfarrar@oucru.org



A swab sample is taken from a pig in the village of Yeh Kuning, Bali, to test for H5N1 influenza.

How to track a flu virus

Four experts pinpoint ways to improve monitoring of H5N1 avian influenza in the field.

YI GUAN

Monitor outbreaks in domestic ducks

Joint Influenza Research Centre, Shantou University Medical College and University of Hong Kong, China

The H5N1 influenza outbreak in Asia is unprecedented: never before has a highly pathogenic avian influenza virus prevailed for so long, spread to so many countries or generated so many genetic variants. Why? Partly because of its persistence in domestic ducks.

In parts of southeast Asia such as China and Vietnam, H5N1 has remained endemic. Elsewhere, small outbreaks last for a very short time. The major difference between these regions is their domestic duck populations — more than 70% of the world's ducks are raised where H5N1 is endemic.

In our 12-year surveillance, more than 65% of the H5N1 viruses my colleagues and I isolated were from domestic ducks. Asymptomatic ducks could shed high concentrations of the virus for several days. Although the H5N1 virus resides in domestic ducks, it can interact with other

subtypes of influenza, for which these birds are part of the natural reservoir. This mixing creates novel variants, which may trigger outbreaks and dissemination of the virus. Domestic ducks probably shelter the H5N1 virus during the summer and then seed the next outbreak, which, in bird populations, usually peaks during the winter.

At present, surveillance of duck populations is limited. Eradication of H5N1 will require more active surveillance in affected areas, along with widespread vaccination of duck populations, segregation of poultry species and local moratoriums on poultry production when outbreaks occur.

RICHARD WEBBY

Improve surveillance of pigs

Department of Infectious Diseases, St Jude Children's Research Hospital, Memphis, Tennessee

One major problem with H5N1 surveillance is the lack of coordinated monitoring in pigs. Although H5N1 is considered to be avian flu, the same mutations that ▶

of the capacity and flexibility needed to respond immediately to the rapidly changing dynamic of infectious diseases such as H5N1, enterovirus 71 or artemisinin-resistant malaria, and can make results available in real time. Such an approach is impossible when the work requires individuals to fly in and out and to analyse samples in another country.

SMALL BUT POWERFUL

There are other great examples of long-term research partnerships between national and international organizations, but they are all too few. These infrastructures are easier to build than many believe — you need only a small group of committed people, a shared vision and ethos, flexible funding that encourages local decision-making, and a focus on excellence. There can be great power in such small institutions — which may need as little as a few hundred thousand dollars a year to operate — if only we made better use of them (G. T. Keusch and C. A. Medlin *Nature* **422**, 561–562; 2003).

Because our research unit is based in the region where the story is unfolding, we can appreciate the social issues that can stymie even the best scientific endeavour. For instance, small-scale backyard poultry farms (often family farms with mixed chickens, ducks and pigs) remain a crucial livelihood and the main source of protein for many households in rural Asia. Because no adequate compensation schemes have been developed to encourage reporting of sick poultry and livestock, the usual responses are to cull all local poultry and apportion blame. Such activities can ruin small farmers and their families.

There is now a window of opportunity to build global scientific capacity before another crisis — such as a new pandemic — hits. This means collaborating with the people who share a vested interest in using the money efficiently and effectively to prevent outbreaks and address daily public-health and clinical issues in their own countries. After living in Vietnam for more than 16 years and raising my family here, I can understand the feeling of urgency. Everyone I work with who sees chickens each day on their way to work, hears about local outbreaks in the news or treats patients is united in the effort to stay one step ahead of H5N1 and other potentially deadly outbreaks. We must share the available knowledge and the tools to make it possible — an undertaking that will require us to shift the centre of gravity for such research to where the needs are greatest. ■

Jeremy Farrar is at the Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. e-mail: jfarrar@oucru.org



A swab sample is taken from a pig in the village of Yeh Kuning, Bali, to test for H5N1 influenza.

How to track a flu virus

Four experts pinpoint ways to improve monitoring of H5N1 avian influenza in the field.

YI GUAN

Monitor outbreaks in domestic ducks

Joint Influenza Research Centre, Shantou University Medical College and University of Hong Kong, China

The H5N1 influenza outbreak in Asia is unprecedented: never before has a highly pathogenic avian influenza virus prevailed for so long, spread to so many countries or generated so many genetic variants. Why? Partly because of its persistence in domestic ducks.

In parts of southeast Asia such as China and Vietnam, H5N1 has remained endemic. Elsewhere, small outbreaks last for a very short time. The major difference between these regions is their domestic duck populations — more than 70% of the world's ducks are raised where H5N1 is endemic.

In our 12-year surveillance, more than 65% of the H5N1 viruses my colleagues and I isolated were from domestic ducks. Asymptomatic ducks could shed high concentrations of the virus for several days. Although the H5N1 virus resides in domestic ducks, it can interact with other

subtypes of influenza, for which these birds are part of the natural reservoir. This mixing creates novel variants, which may trigger outbreaks and dissemination of the virus. Domestic ducks probably shelter the H5N1 virus during the summer and then seed the next outbreak, which, in bird populations, usually peaks during the winter.

At present, surveillance of duck populations is limited. Eradication of H5N1 will require more active surveillance in affected areas, along with widespread vaccination of duck populations, segregation of poultry species and local moratoriums on poultry production when outbreaks occur.

RICHARD WEBBY

Improve surveillance of pigs

Department of Infectious Diseases, St Jude Children's Research Hospital, Memphis, Tennessee

One major problem with H5N1 surveillance is the lack of coordinated monitoring in pigs. Although H5N1 is considered to be avian flu, the same mutations that ▶

▶ allow the virus to be transmitted between ferrets could make it more contagious among both humans and pigs. Any virus that circulates in pigs can be a risk to humans — just look at the furore over the 2009 H1N1 pandemic, which probably originated in pigs. And, because pigs can harbour multiple strains of the influenza virus, they are good incubators for mutants — including those that might make H5N1, for example, more contagious.

In most countries where pigs are farmed, one only has to look for swine influenza viruses to find them. Systematic surveillance in pigs is conducted in some countries, but in other parts of the world there is none. We must encourage all countries with large pig populations to perform systematic surveillance and to report what they find.

My colleagues and I have conducted surveillance of US farms for the past few years, sampling healthy pigs once a month to get a baseline of influenza activity so that we can act quickly at the first hints of unusual activity. We've found that 2–4% of all seemingly healthy animals harbour some type of influenza virus.

We, along with others in the swine-influenza group of the OIE–FAO Network of Expertise on Animal Influenza (OFFLU),

deposit our data in a public database. Swine-influenza surveillance done now is scattered — among universities, industry and government agencies. A good first step, as the OFFLU group urges, is to begin to coordinate that activity with measures such as the creation of a centralized database.

ILARIA CAPUA

Restore ties lost in the Arab Spring

Director, International Reference Laboratory for Avian Influenza and Newcastle Disease, Veterinary Public Health Institute, Legnaro, Italy

Over the past six years, my lab has developed collaborations with veterinary authorities in several African and Middle Eastern countries to conduct surveillance of avian influenza. The Arab Spring has transformed society in some of these countries, but it has negatively influenced our efforts to monitor H5N1 in nature.

In Egypt, outbreaks in poultry are

widespread, and 58 people have so far died of the disease. The country had begun developing a system to combat H5N1 with the aid of international agencies and donors. The system was not perfect, but it was the result of years of work to build capacity and train personnel. Since the social unrest began, things have changed. Some of our local contacts have moved to other positions, leaving gaps in expertise.

At present, we are working mainly with local universities and industry. On average, we can analyse viral genomes within 4–6 months after they are isolated from the field — a delay that could easily nullify the benefits of monitoring a potentially pandemic virus while it is still in the animal reservoir.

We recently met with newly appointed veterinary officials in Egypt and discussed collaboration and ways forward. We are confident that improved surveillance will be one of the positive outcomes of these discussions. The support of international organizations is crucial to the success of such efforts.

JONAS WALDENSTRÖM

Learn more about the role of wild birds

School of Natural Sciences, Linnaeus University, Kalmar, Sweden

Migrating birds can be infected with H5N1 and could potentially spread the virus along migratory flyways. Despite intense, active surveillance, we still know frustratingly little about how H5N1 is transmitted in the environment and between wild and domestic birds.

My colleagues and I have screened samples from 30,000 European waterfowl without finding H5N1 — or any other highly pathogenic influenza virus — during or after the H5N1 outbreak in this population. Haphazard sampling and poor logistics may contribute to poor detection rates, or the virus simply may not be in the population when we search for it.

We must provide greater resource support for areas at high risk of H5N1 and develop better tools for detecting the virus in wild birds so that we can better assess the role of these populations in virus perpetuation, on local and global scales. Surveillance programmes should be complemented with more-targeted studies that address key questions in H5N1 epidemiology, such as how virus transmission occurs at the interface between domestic poultry and wild birds.



A man sells fowl at a market in Cairo. In Egypt, H5N1 was first observed in 2006 in domestic poultry.

L. YIGUANG/XINHUA/WFPN/PHOTOSHOT

Earth's weather and climate are influenced by variations in its orbit as well as by oscillations in its internal systems.

CLIMATE SCIENCE

A delicate balance

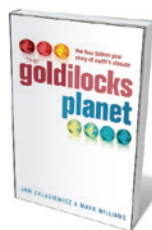
Earth's climate and biosphere have always shaped one another. **James F. Kasting** approves of an attempt to reveal the planet's future by reading its past.

Earth has warmed rapidly before. About 55 million years ago, the temperature of the planet rose by as much as 8 °C over 20,000 years and remained elevated for roughly 100,000 years. The cause is unknown, but it may have been a result of cage-like methane clathrate molecules in the sea floor destabilizing and releasing into the atmosphere huge amounts of greenhouse gas. This Palaeocene–Eocene Thermal Maximum is of great interest to climatologists: the estimated temperature increase is similar to the future warming predicted owing to human activities, although we are perturbing the climate system much faster.

Such events show that if you want to understand the climate's future, you need to learn about its past. *The Goldilocks Planet* — named after the concept that Earth, unlike its planetary neighbours, is just right for life, neither too hot nor too cold — describes that past, from its fiery beginnings to a warmed world. The book is stronger on geologically recent history than on the deep past, but it shows that the story of past climate change is a powerful way to convey the realities and risks of human-induced global warming.

Jan Zalasiewicz and Mark Williams are both experts in Quaternary micropalaeontology: they study the microbial fossil record from just under 2.6 million years ago to the present. They are also well versed in isotopic biogeochemistry, deciphering the planet's history from the chemical traces of life.

The climate of the Quaternary period has been defined by a cycle of successive ice ages and interglacial periods. The authors discuss in depth what drives this, describing Milankovitch cycles — climate excursions caused by variations in Earth's orbit — and the less well-known Dansgaard–Oeschger and Heinrich events, both climate variations on shorter time scales, thought to result from oscillations within Earth's system. These sections form



The Goldilocks Planet: The 4 Billion Year Story of Earth's Climate
JAN ZALASIEWICZ AND MARK WILLIAMS
Oxford University Press: 2012. 272 pp.
£16.99, \$29.95

a good introduction to the topic for non-specialists.

There are some amusing stories here. The authors tell of a journey through the Drake Passage, which runs between South America and Antarctica, in the UK Royal Navy's ice-breaker *HMS Endurance*: a flat-bottomed vessel that “rolls like a pig” in high seas. We learn about the geologist Nicholas Shackleton's love for clarinets, and what boron isotopes reveal about the acidity of ancient oceans.

There is a fascinating account of the closing of the Isthmus of Panama, some 3 million years ago. This allowed land animals such as armadillos — introduced in the book as Texan roadkill — to migrate between North and South America, and increased the salinity gradient between the Atlantic and Pacific oceans, helping to establish the modern pattern of thermohaline ocean circulation. The authors are true experts in this field.

Zalasiewicz and Williams are also knowledgeable about the climatic history of the Phanerozoic eon, the time from 542 million years ago to the present in which there is a good fossil record of multicellular plants and animals. From tales of traipsing around ▶

ISS/NASA

▶ England's rocky shorelines, I learned how graptolites — microscopic colonial animals that lived in sediments during the early Palaeozoic era, about 450 million years ago — vanished from the fossil record when the climate cooled, because the cold, oxygen-rich water that penetrated the deep sea let predators invade the sediments and eat them.

But the book's coverage of climate evolution during the earliest nine-tenths of Earth's history — the Precambrian era, 542 million years ago and earlier — is neither so detailed nor so scientifically balanced. One chapter describes most of this time interval, and another focuses on the Late Proterozoic Snowball Earth glaciations, when ice repeatedly covered nearly all Earth's surface, around 635 million years ago.

Zalasiewicz and Williams give scant attention to the carbonate-silicate cycle that many Earth scientists believe is the key to Earth's Goldilocks status. The basic idea is that volcanoes add carbon to the atmosphere and the sea, and the weathering of silicate minerals on the continents and the deposition of carbonate sediments in the oceans take it out. Weathering slows as the climate cools, so carbon dioxide builds up, warming the climate and creating negative feedback. This feedback is mentioned as causing the Snowball Earth to melt, but its importance in regulating climate in general is not really discussed.

Neither are the authors the best guides to Snowball Earth events. They give too much weight to discarded ideas such as the high-obliquity hypothesis, which argues that the glaciations resulted from the tilt of the Earth's axis, and they omit to mention the latest thinking and evidence.

The last chapter of *The Goldilocks Planet* deals with the Anthropocene epoch — a term popularized by Nobel-prizewinning atmospheric chemist Paul Crutzen to describe the geological epoch in which humans have significantly modified the Earth's climate. This is well-trodden ground, but the discussion is on the mark, and the preceding review of climate history gives it credibility. If Earth's climate is as sensitive as it seems to be, then how could it not respond to the massive greenhouse forcing that humans would create by burning a significant fraction of the available fossil fuels?

Pennsylvania State University teaches a general-education Earth science course that approaches global warming in the same way: reviewing climate history to give a context for the anticipated future. It works well for us, and it works for *The Goldilocks Planet*, too. ■

James F. Kasting is a distinguished professor of geosciences at Pennsylvania State University in University Park, and the author of *How to Find a Habitable Planet* (Princeton Univ. Press, 2009). e-mail: kasting@essc.psu.edu



Ferdinand de Saussure was hugely influential in the social sciences, despite publishing little.

LINGUISTICS

Sound sculptor

John A. Goldsmith is intrigued by the life of a linguistics giant who felt himself to be a failure.

“Any life when viewed from the inside is simply a series of defeats,” wrote George Orwell. Each serious study of a great scientist's life is bound to leave us reflecting on that truth, and linguist John E. Joseph's monumental *Saussure* is no exception.

Ferdinand de Saussure (1857–1913) was one of the great nineteenth-century linguists, and Joseph's book, the first comprehensive biography, sheds brilliant light on his life and work. This rich account — sympathetic, respectful and sensitive to political and intellectual context — reveals how Saussure, a dazzling and driven scholar from a bourgeois Swiss family, blazed trails to new vistas of social science that opened out in the century after his death.

Saussure emerges as a complex individual. As Joseph shows us, his virtuosity was

Saussure
JOHN E. JOSEPH
Oxford University
Press: 2012. 800 pp.
£30.00, \$55.00

counterbalanced by a series of unfortunate failures to get projects finished. That Orwellian ‘inside story’ hums away in

the background of the book.

Two works of genius bookended Saussure's life. The first was a revolutionary monograph on Proto-Indo-European — the ‘raw material’ of modern languages from English to Sanskrit — self-published when he was just 21. Saussure's contribution was to deduce that ancient Indo-European must have contained certain sounds that had disappeared from more modern languages and so were undetectable in linguistic history. The prediction was confirmed much later in an analysis of Hittite documents from the thirteenth century BC.

AKG-IMAGES

► England's rocky shorelines, I learned how graptolites — microscopic colonial animals that lived in sediments during the early Palaeozoic era, about 450 million years ago — vanished from the fossil record when the climate cooled, because the cold, oxygen-rich water that penetrated the deep sea let predators invade the sediments and eat them.

But the book's coverage of climate evolution during the earliest nine-tenths of Earth's history — the Precambrian era, 542 million years ago and earlier — is neither so detailed nor so scientifically balanced. One chapter describes most of this time interval, and another focuses on the Late Proterozoic Snowball Earth glaciations, when ice repeatedly covered nearly all Earth's surface, around 635 million years ago.

Zalasiewicz and Williams give scant attention to the carbonate-silicate cycle that many Earth scientists believe is the key to Earth's Goldilocks status. The basic idea is that volcanoes add carbon to the atmosphere and the sea, and the weathering of silicate minerals on the continents and the deposition of carbonate sediments in the oceans take it out. Weathering slows as the climate cools, so carbon dioxide builds up, warming the climate and creating negative feedback. This feedback is mentioned as causing the Snowball Earth to melt, but its importance in regulating climate in general is not really discussed.

Neither are the authors the best guides to Snowball Earth events. They give too much weight to discarded ideas such as the high-obliquity hypothesis, which argues that the glaciations resulted from the tilt of the Earth's axis, and they omit to mention the latest thinking and evidence.

The last chapter of *The Goldilocks Planet* deals with the Anthropocene epoch — a term popularized by Nobel-prizewinning atmospheric chemist Paul Crutzen to describe the geological epoch in which humans have significantly modified the Earth's climate. This is well-trodden ground, but the discussion is on the mark, and the preceding review of climate history gives it credibility. If Earth's climate is as sensitive as it seems to be, then how could it not respond to the massive greenhouse forcing that humans would create by burning a significant fraction of the available fossil fuels?

Pennsylvania State University teaches a general-education Earth science course that approaches global warming in the same way: reviewing climate history to give a context for the anticipated future. It works well for us, and it works for *The Goldilocks Planet*, too. ■

James F. Kasting is a distinguished professor of geosciences at Pennsylvania State University in University Park, and the author of *How to Find a Habitable Planet* (Princeton Univ. Press, 2009). e-mail: kasting@essc.psu.edu



Ferdinand de Saussure was hugely influential in the social sciences, despite publishing little.

LINGUISTICS

Sound sculptor

John A. Goldsmith is intrigued by the life of a linguistics giant who felt himself to be a failure.

“Any life when viewed from the inside is simply a series of defeats,” wrote George Orwell. Each serious study of a great scientist's life is bound to leave us reflecting on that truth, and linguist John E. Joseph's monumental *Saussure* is no exception.

Ferdinand de Saussure (1857–1913) was one of the great nineteenth-century linguists, and Joseph's book, the first comprehensive biography, sheds brilliant light on his life and work. This rich account — sympathetic, respectful and sensitive to political and intellectual context — reveals how Saussure, a dazzling and driven scholar from a bourgeois Swiss family, blazed trails to new vistas of social science that opened out in the century after his death.

Saussure emerges as a complex individual. As Joseph shows us, his virtuosity was

Saussure
JOHN E. JOSEPH
Oxford University
Press: 2012. 800 pp.
£30.00, \$55.00

counterbalanced by a series of unfortunate failures to get projects finished. That Orwellian ‘inside story’ hums away in

the background of the book.

Two works of genius bookended Saussure's life. The first was a revolutionary monograph on Proto-Indo-European — the ‘raw material’ of modern languages from English to Sanskrit — self-published when he was just 21. Saussure's contribution was to deduce that ancient Indo-European must have contained certain sounds that had disappeared from more modern languages and so were undetectable in linguistic history. The prediction was confirmed much later in an analysis of Hittite documents from the thirteenth century BC.

AKG-IMAGES

The second work was a series of groundbreaking lectures that Saussure gave towards the end of his life, and which was later transformed, through notes taken down and edited by his students, into the great *Course on General Linguistics*, published in 1916. Much of Saussure's fame came from a book that he never wrote.

Modern linguists remember the wunderkind who followed his stunning early essay with a doctoral dissertation on Sanskrit just a year later. Saussure is also renowned as a key figure in the rise of structuralism, a method that profoundly influenced the social sciences by looking for a universal structure behind human behaviour and social activity. In linguistics, the method can be used to analyse the evolution of language by searching for patterns and symmetries in sounds.

But there were, as Joseph clearly recounts, setbacks. He shows how, after the publication of the monograph, Saussure's teachers at the University of Leipzig in Germany were stunned. They saw his dazzling, innovative reconstruction of Indo-European vowels as a brazen elaboration on what he had been taught in their courses. Fortunately, his doctoral dissertation was much more conventional, and he was granted a degree. When he left Germany, he was not much missed by his teachers.

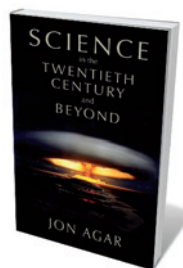
Saussure then spent a decade teaching in Paris, publishing nothing except a few brief notes and unable to obtain a professorship. He spent a good deal of money he didn't have betting at the races, making up for his losses by winning at poker. Saussure eventually returned to Switzerland, where he was named a professor at the University of Geneva. He continued to teach, but never had many students.

Although Saussure filled scores of notebooks with his research, he never succeeded in producing a book that satisfied his own standards. His students produced a publication in his honour on his 50th birthday, but he died a few years later, unable to see any accomplishments in his life past what he did as a very young man. Yet, as Joseph shows, the *Course on General Linguistics* continues to have enormous influence on thinkers not only in linguistics, but also in anthropology, sociology and literary criticism.

Weighing up the truth about the whole person — life and work — is no easy thing. Saussure was never to have the satisfaction of understanding the vast reach of his own work. And although we do at least have that, we must sift through the evidence to make our estimation of the life and the man. ■

John A. Goldsmith is the Edward Carson Waller Distinguished Service Professor of Linguistics and Computer Science at the University of Chicago.
e-mail: goldsmith@uchicago.edu

Books in brief



Science in the Twentieth Century and Beyond

Jon Agar POLITY 256 pp. £30 (2012)

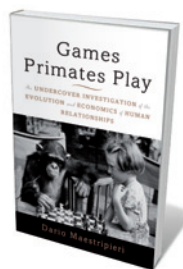
Global in scope and fresh in approach, this monumental history lays out the evolution of science during a tumultuous century. Philosopher of science Jon Agar casts research as a way of solving problems generated by human activity in arenas such as health, warfare, civil administration and agriculture. Starting with the new physics and the breakthroughs of figures from James Clerk Maxwell to Albert Einstein, he travels through the life sciences, psychology, the maelstrom of science in the two world wars, the atomic age, upheavals of the 1960s and current environmental challenges.



Innovation Generation: How to Produce Creative and Useful Scientific Ideas

Roberta Ness OXFORD UNIVERSITY PRESS 272 pp. \$29.95 (2012)

Obesity, Alzheimer's disease, cancer and climate change are very much with us, says Roberta Ness, yet innovation that could mitigate them has slowed catastrophically in US science. Ness, vice-president for innovation at the University of Texas Health Science Center at Houston, outlines a method to ignite creativity. Through role models and exercises, she shows how better metaphors and observation can shift paradigms; and how specific issues can be solved with the right questions, the right analogies and group intelligence.



Games Primates Play: An Undercover Investigation of the Evolution and Economics of Human Relationships

Dario Maestri BASIC BOOKS 320 pp. \$27.99 (2012)

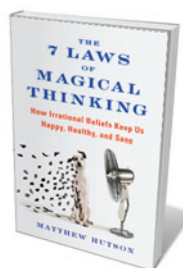
Evolutionary biologist Dario Maestri uncovers the roots of human social behaviour using psychology, behavioural science and economics. Reasoning that social selective pressures are similar in humans and other primates — and roping in 'rational' models such as game theory — he examines everyday situations from multiple perspectives. Whether scoping out the 'elevator dilemma' of sharing a confined space with strangers, the human tendency to nepotism or the "economics of love", Maestri argues his case compellingly.



Net Smart: How to Thrive Online

Howard Rheingold MIT PRESS 272 pp. \$24.95 (2012)

Fragmented attention, aimless dabbling — the pitfalls of Internet misuse are well known. Social-media writer Howard Rheingold argues that the solution to "always on" media is mindfulness and cooperation. His recipe for digital literacy, based on 30 years of Internet immersion, is to hone attention, participation skills, critical approaches to information, collaboration and "network smarts". Rheingold's observations and solutions — from how tweeting is fuelled by dopamine to how to craft a thoughtful network — are informed by science and illustrated with apt, entertaining anecdotes.



The 7 Laws of Magical Thinking: How Irrational Beliefs Keep Us Happy, Healthy, and Sane

Matthew Hutson HUDSON STREET 304 pp. \$25.95 (2012)

Irrationality, says science writer Matthew Hutson, is universal, and is essential to the way humans function. Uniting findings from neuroscience, cognitive science and evolution, he argues that magical thinking gives us crucial feelings of connectedness, control and meaning. Hutson analyses the call of the numinous in a range of beliefs: the 'sacred' essence in wedding rings or signed footballs, lucky numbers, an afterlife, fate, psychic powers and more.

The second work was a series of groundbreaking lectures that Saussure gave towards the end of his life, and which was later transformed, through notes taken down and edited by his students, into the great *Course on General Linguistics*, published in 1916. Much of Saussure's fame came from a book that he never wrote.

Modern linguists remember the wunderkind who followed his stunning early essay with a doctoral dissertation on Sanskrit just a year later. Saussure is also renowned as a key figure in the rise of structuralism, a method that profoundly influenced the social sciences by looking for a universal structure behind human behaviour and social activity. In linguistics, the method can be used to analyse the evolution of language by searching for patterns and symmetries in sounds.

But there were, as Joseph clearly recounts, setbacks. He shows how, after the publication of the monograph, Saussure's teachers at the University of Leipzig in Germany were stunned. They saw his dazzling, innovative reconstruction of Indo-European vowels as a brazen elaboration on what he had been taught in their courses. Fortunately, his doctoral dissertation was much more conventional, and he was granted a degree. When he left Germany, he was not much missed by his teachers.

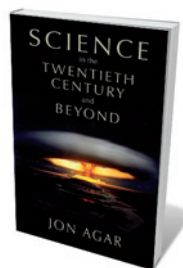
Saussure then spent a decade teaching in Paris, publishing nothing except a few brief notes and unable to obtain a professorship. He spent a good deal of money he didn't have betting at the races, making up for his losses by winning at poker. Saussure eventually returned to Switzerland, where he was named a professor at the University of Geneva. He continued to teach, but never had many students.

Although Saussure filled scores of notebooks with his research, he never succeeded in producing a book that satisfied his own standards. His students produced a publication in his honour on his 50th birthday, but he died a few years later, unable to see any accomplishments in his life past what he did as a very young man. Yet, as Joseph shows, the *Course on General Linguistics* continues to have enormous influence on thinkers not only in linguistics, but also in anthropology, sociology and literary criticism.

Weighing up the truth about the whole person — life and work — is no easy thing. Saussure was never to have the satisfaction of understanding the vast reach of his own work. And although we do at least have that, we must sift through the evidence to make our estimation of the life and the man. ■

John A. Goldsmith is the Edward Carson Waller Distinguished Service Professor of Linguistics and Computer Science at the University of Chicago.
e-mail: goldsmith@uchicago.edu

Books in brief



Science in the Twentieth Century and Beyond

Jon Agar POLITY 256 pp. £30 (2012)

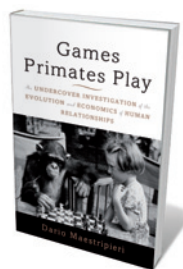
Global in scope and fresh in approach, this monumental history lays out the evolution of science during a tumultuous century. Philosopher of science Jon Agar casts research as a way of solving problems generated by human activity in arenas such as health, warfare, civil administration and agriculture. Starting with the new physics and the breakthroughs of figures from James Clerk Maxwell to Albert Einstein, he travels through the life sciences, psychology, the maelstrom of science in the two world wars, the atomic age, upheavals of the 1960s and current environmental challenges.



Innovation Generation: How to Produce Creative and Useful Scientific Ideas

Roberta Ness OXFORD UNIVERSITY PRESS 272 pp. \$29.95 (2012)

Obesity, Alzheimer's disease, cancer and climate change are very much with us, says Roberta Ness, yet innovation that could mitigate them has slowed catastrophically in US science. Ness, vice-president for innovation at the University of Texas Health Science Center at Houston, outlines a method to ignite creativity. Through role models and exercises, she shows how better metaphors and observation can shift paradigms; and how specific issues can be solved with the right questions, the right analogies and group intelligence.



Games Primates Play: An Undercover Investigation of the Evolution and Economics of Human Relationships

Dario Maestri BASIC BOOKS 320 pp. \$27.99 (2012)

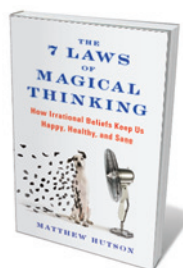
Evolutionary biologist Dario Maestri uncovers the roots of human social behaviour using psychology, behavioural science and economics. Reasoning that social selective pressures are similar in humans and other primates — and roping in 'rational' models such as game theory — he examines everyday situations from multiple perspectives. Whether scoping out the 'elevator dilemma' of sharing a confined space with strangers, the human tendency to nepotism or the "economics of love", Maestri argues his case compellingly.



Net Smart: How to Thrive Online

Howard Rheingold MIT PRESS 272 pp. \$24.95 (2012)

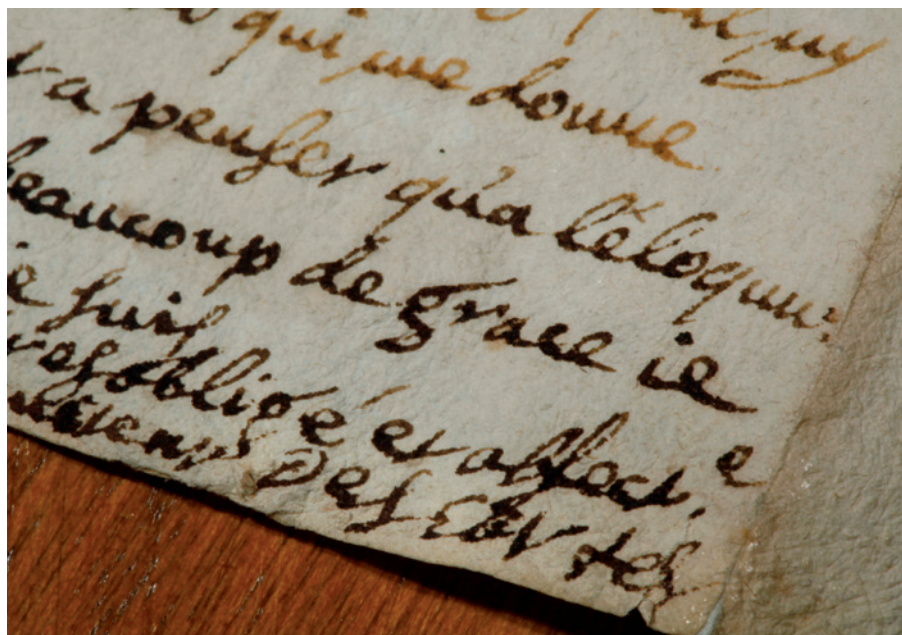
Fragmented attention, aimless dabbling — the pitfalls of Internet misuse are well known. Social-media writer Howard Rheingold argues that the solution to "always on" media is mindfulness and cooperation. His recipe for digital literacy, based on 30 years of Internet immersion, is to hone attention, participation skills, critical approaches to information, collaboration and "network smarts". Rheingold's observations and solutions — from how tweeting is fuelled by dopamine to how to craft a thoughtful network — are informed by science and illustrated with apt, entertaining anecdotes.



The 7 Laws of Magical Thinking: How Irrational Beliefs Keep Us Happy, Healthy, and Sane

Matthew Hutson HUDSON STREET 304 pp. \$25.95 (2012)

Irrationality, says science writer Matthew Hutson, is universal, and is essential to the way humans function. Uniting findings from neuroscience, cognitive science and evolution, he argues that magical thinking gives us crucial feelings of connectedness, control and meaning. Hutson analyses the call of the numinous in a range of beliefs: the 'sacred' essence in wedding rings or signed footballs, lucky numbers, an afterlife, fate, psychic powers and more.



Dozens of René Descartes' letters have been missing since the mid-nineteenth century.

Q&A Erik-Jan Bos

Descartes' decipherer

Erik-Jan Bos, a philosopher at Utrecht University in the Netherlands, unearthed research gold with an Internet search. In putting together a critical edition of René Descartes' correspondence, due out in 2014, he discovered a stolen, never-before-published letter from the seventeenth-century French philosopher and mathematician. In the run-up to Descartes' 416th birthday on 31 March, Bos discusses the hazards of chasing him down.

Tell us about your Google search for a stolen letter.

I searched for 'Descartes' and 'autograph letter', and got a hit at Haverford College in Pennsylvania on the first page. The listing immediately caught my attention. I had been using those search terms for a few years, so I knew the first 30 hits very well. From this letter, to French mathematician Marin Mersenne, we learn that Descartes had changed the introduction to his 1641 book *Meditations on First Philosophy* at the last minute. The letter had been stolen in the nineteenth century by Guglielmo Libri, a gifted historian of mathematics and an ardent bibliophile and collector. Eventually he not only bought manuscripts, but also stole them. He became an inspector of French public libraries, and looted them. He took about 80 autographed letters of Descartes from the Institut de France in Paris.

How did they end up in the United States?

After he had cherished them for several years, Libri decided to sell. That is how his collection got dispersed. The scandal became public in 1848 and he fled to England. The majority of the letters returned



to France, but about 30 were untraceable. Some were completely unknown and never published, such as the one I discovered at Haverford. It was a very rare find.

Have any letters proven problematic?

The standard edition of Descartes' correspondence contains one letter that is a fake. It was copied from a novel of the late seventeenth century, which ridicules the belief that people can talk to spirits. The most notorious forger is Denis Vrain-Lucas, who sold thousands of counterfeited letters, supposedly by Descartes, Jesus, Mary Magdalene, Charlemagne, Aristotle, Lazarus before and after his resurrection, and so on. Once unmasked, Vrain-Lucas pleaded innocent, saying that the question of authenticity is trivial. The problem is that there is money to be made from forgery. A genuine letter from Descartes could sell for US\$200,000 or more.

Why create a new edition of correspondence?

The standard edition is a century old and

contains a lot of supplements and errata; the next edition contains even more. The corpus is a nightmare. If you want to find particular material, you will get lost. With a new edition we can check formulae and mathematical texts. In previous editions, complete lines were left out, making the maths incomprehensible at times. Claude Clerselier, the first editor of Descartes' correspondence in the seventeenth century, wrote in the preface to the first collection that many of the manuscripts were difficult to decipher, so he had to guess. We can't be sure exactly what he did because the manuscripts he edited have vanished. After he died they went to a French scholar who died, then on to the next scholar who died, and finally all the material went back to the mother of the first scholar. I don't know what she did — maybe put them on the fire during a cold winter?

How has the work of finding correspondence changed?

The amount of labour my predecessors had to go through is unimaginable. After I Googled the lost letter, the librarian at Haverford immediately took digital pictures of the manuscript, plugged in the camera and sent them to me. Twenty years ago, I would have had to wait for weeks. As for the search, half a century ago, if you went to a library in Paris, you could search for weeks and still miss important material. In the past decade, inventories of manuscripts have come online. But you have to be clever. If you just search for 'Descartes', you get millions and millions of hits.

Why study Descartes' letters?

He comes alive. He looked down on contemporary philosophers, scientists and mathematicians, including Pierre de Fermat, with utter disdain. He also thought that he was always right. And there are unexpectedly personal letters. We have one by Descartes written to the local bailiff pleading for leniency towards an accused murderer. He wrote wonderful letters to friends who had lost loved ones. Other highlights include a marriage contract drawn up in 1644, for which one of the witnesses was Descartes. The bride was the mother of Descartes' daughter. He recognized his fatherhood but never married the mother, although presumably he saw to it that she got married. That contract was a way to take care of her.

Has reading Descartes helped you with maths?

An encounter with maths in his letters can be like reading Greek. The seventeenth-century way of doing maths was different from today's: they were interested in other problems. We live in a post-Cartesian era, mathematically and philosophically. ■

INTERVIEW BY STEPHEN ORNES

Correspondence

Peak oil is affecting the economy already

James Murray and David King sound the alarm in pointing out that oil's tipping point has passed (*Nature* **481**, 433–435; 2012). The days of being able to produce oil cheaply and easily are over, and the economic effects are upon us. We believe that the 'peak oil' issue is as important as climate change, and more urgent. We call for peak oil to be considered more seriously as a subject of peer-reviewed research.

Economic growth began to stall at around the same time as conventional oil production in 2005. Oil-exporting nations are consuming more of their own output every year, reducing availability for the rest of the world. Yet importers who rely on oil exports include some of the world's largest economies.

The energy return on investment for oil is declining globally. There is now a growing body of research exploring the connections between financial and energy returns on oil investment, and economic health as a whole.

David J. Murphy* *Northern Illinois University, DeKalb, USA.*
djmurphy@niu.edu

**On behalf of 8 co-signatories (see go.nature.com/qlcack for a full list).*

Turing: keep his work in perspective

In a cascade of plaudits, George Dyson, Sydney Brenner and Barry Cooper each suggest that Alan Turing's bridging of logic and machines laid the foundation for digital computers, built in the 1940s under John von Neumann (*Nature* **482**, 459–460, 461 and 465; 2012).

Turing and von Neumann are both heroes of mine. But, in essence, Turing's famous 1936 paper on incomputability was merely an elegant rephrasing and reuse of mathematician Kurt Gödel's 1931 results and

techniques. Gödel devised a more cumbersome, integer-based language to describe a universal algorithmic theorem-prover, which allowed him to identify the fundamental limits of mathematics and provability.

Neither did Turing's paper have any impact on the construction of the first program-controlled universal (and digital) computer: that was built in Berlin by Konrad Zuse in 1935–41, at least 3 years before anyone else's. Zuse's 1936 patent application mentioned an architecture like that of von Neumann's (which von Neumann described in 1945), with programs and data that could be modified in storage. Computing firm IBM was well aware of these breakthroughs and funded Zuse's 1946 start-up through an option on his patents.

Jürgen Schmidhuber *Dalle Molle Institute for Artificial Intelligence (IDSIA), University of Lugano; and SUPSI, Manno-Lugano, Switzerland.*
juergen@idsia.ch

Turing: a formal clash of codes

Sydney Brenner argues that cells and living organisms are good examples of Turing and von Neumann machines (*Nature* **482**, 461; 2012). But the nature of living matter cannot be properly accommodated within such a theoretical framework.

This is because the language that codes machine programs is not compatible with that of the genetic code. Languages controlling Turing and von Neumann machines are formal algorithms, in which syntax determines meaning independently of context. Gene expression depends on environmental context, however, so cannot be similarly treated as a formal language.

Syntax in DNA may convey different and even contradictory meanings, depending on the cellular agents that exploit the

coded information according to their respective situations.

Guenther Witzany *Telos-Philosophische Praxis, Bürmoos, Austria.*

František Baluška *Institute of Cellular and Molecular Botany, University of Bonn, Germany.*
baluska@uni-bonn.de

The case for brain imaging technology

Olivier Oullier questions the commercial and judicial use of brain-scan technology to predict or judge human behaviour (*Nature* **483**, 7; 2012). His arguments undermine a major driver of academic funding and research — its potential for commercial application. Moreover, banning the commercial use of neuroimaging could encourage government interference in the development of a promising and widely applicable tool.

Society demands benefits in return for investment in scientific research. Universities are expected to be entrepreneurial and to collaborate with industry, promote spin-offs and capitalize on intellectual property.

An emerging technology should be judged on its specificity, sensitivity and predictive value. Neuroimaging results must be assessed in their specific context, and not broadly dismissed. Functional magnetic resonance imaging used to measure motivation in response to health campaigns, for example, was found to have greater predictive value than self-reported intentions (E. B. Falk *et al.* *J. Neurosci.* **30**, 8421–8424; 2010).

The technique is also able to identify paedophilic inclinations with 95% accuracy (J. Ponseti *et al.* *Arch. Gen. Psychiatry* **69**, 187–194; 2012). Careful investigation will establish whether or not such findings can one day be used as evidence in court.

Bernd Weber* *Center for Economics and Neuroscience,*

University of Bonn; and Department of Epileptology, University Hospital of Bonn, Germany.

bernd.weber@ukb.uni-bonn.de

**On behalf of 13 co-authors; competing financial interests declared (go.nature.com/9dh9e2).*

True value of climate fund's contribution

Edward Barbier grossly underestimates the Adaptation Fund's contribution to sustainable development (*Nature* **483**, 30; 2012). The fund has so far received US\$273 million and committed \$109 million (see go.nature.com/sgkgr8), not \$12.6 million as quoted.

The Adaptation Fund is supported by a carbon market, one of the innovative revenue models Barbier mentions. Under the Clean Development Mechanism of the United Nations Framework Convention on Climate Change, developed countries can trade carbon credits generated by sponsoring emissions-reduction projects in developing countries, helping them to meet their emissions targets under the Kyoto Protocol.

Of the revenue generated by these transactions, 2% goes to the Adaptation Fund for programmes to help developing countries adapt to climate change. These tackle such problems as coastal erosion from sea-level rise, floods from glacial melt and drought-induced falls in crop yields. By January 2012, the fund had spent \$30.14 million on such projects.

The carbon market is shrinking because of the global financial crisis and uncertainties over the scale of demand for emissions-reduction credits. However, the Adaptation Fund's structure means it can also accept funds from governments, the private sector and individuals.

Marcia Levaggi *Adaptation Fund Board Secretariat, Global Environment Facility, Washington DC, USA.*
mlevaggi@thegef.org

Wylie Walker Vale Jr

(1941–2012)

Endocrinologist who deduced the molecular structure of stress hormones.

Wylie Vale and his colleagues answered a long-standing biological riddle: which substance controls the body's 'fight or flight' response? Vale's isolation of corticotropin-releasing factor (CRF) in 1981 and his exploration of the molecular basis of stress hormones opened the way to drugs for treating hypertension, heart disease, obesity, diabetes and depression.

With his passing at the age of 70, the worlds of physiology, neuroscience and peptide biology lost a charismatic leader. Vale effervesced with energy and curiosity, and loved big problems. He believed that life was neither fate nor serendipity, and viewed one's choice of friends, spouse, mentors and colleagues as the most important elements. Born and raised in Houston, Texas, Vale said that marrying his high-school sweetheart Betty was his first and best decision, and attending Houston's Rice University his next.

After his graduation in 1964, Vale pursued his PhD in physiology at Baylor University College of Medicine, also in Houston. His thesis work, completed in 1968, set him on a lifelong quest to dissect the neuroendocrine basis of physiology and behaviour. Vale's sense of adventure was piqued at Baylor by physiologist Roger Guillemin's new science of neuroendocrinology. In the late 1940s at the University of Montreal in Quebec, Canada, Guillemin had teamed up with Viennese physician Hans Selye, who proposed that stress was a specific biological phenomenon and had a neuroendocrine basis.

In 1970, Vale and the Guillemin group moved to the Salk Institute in La Jolla, California. Around this time, with CRF proving too difficult to work on, they concentrated on characterizing other hormones. They reported the structure of the first hypothalamic peptide, thyrotropin-releasing hormone — comprising only three amino acids — highlighted in a landmark *Nature* paper in 1970. And in 1972 and 1973, the group described the factors controlling the release of follicle-stimulating hormone, luteinizing hormone and growth hormone — all key to human development. These results led to the 1977 Nobel Prize in

Physiology or Medicine for Guillemin and Andrew Schally. Vale was in the audience at the Stockholm ceremony, and Guillemin said several times that he should have been on the stage with them.

Yet CRF, the pinnacle of hypothalamic theory, remained shrouded in mystery. The nature of CRF had been debated for dec-

a game-changing expedition to isolate CRF and reveal its molecular structure.

When I arrived at the Salk Institute that year to study steroid and thyroid signalling mechanisms, I was attracted by luminaries such as Robert Holley, Francis Crick, Renato Dulbecco and Guillemin. But I soon found camaraderie in a motley crew of young scientists who met every Saturday morning for tennis and coffee — a tradition that lives on today. At that time, our group included Vale, virologist Inder Verma, and neurobiologists Jean Rivier and Stephen Heinemann, with the occasional guest appearance by tyrosine-kinase expert Tony Hunter. We became lifelong friends, collaborators and colluders, touring the world together in the frantic way that scientists do.

From the outset, it was clear that Vale was on a mission, betting his and his colleagues' careers on a risky gambit — one that in just three years yielded a picture of the crucial 41-amino-acid peptide and a view on to the vast landscape of the stress response it controls. Vale's 1981 breakthrough paper remains in the stratosphere of high-impact studies, and serves as a testimony to his vision and determination. Vale also parlayed his discoveries into two biotechnology companies, founding Neurocrine Biosciences in San Diego, California, and Acceleron Pharma in Cambridge, Massachusetts, to translate these findings into new drugs.

At 70 years old, Wylie was energetic and not ready to retire. He loved adventuring, hiking, slack-key guitar music and sharing a meal and an excellent bottle of wine with an irreverent group of co-conspirators while puzzling life's mysteries. After a lovely evening with friends at his home in Hana, Hawaii, he went to bed and never woke up. Wylie knew that no one was immune from life's slings and arrows, but he enjoyed great fortune and would not have begrudged even his own passing. ■

Ronald Evans is an Investigator of the Howard Hughes Medical Institute and the March of Dimes Chair in Molecular and Developmental Biology at the Salk Institute, La Jolla, California 92037, USA.
e-mail: evans@salk.edu



SALK INST. FOR BIOLOGICAL STUDIES

ades, heightening its allure. Thought to be locked within the hypothalamus, it remained elusive because only tiny amounts could be extracted.

Vale was undeterred. In 1978, he announced to Guillemin in a handwritten letter that he would be leaving the lab and starting his own group to tackle CRF. He set up base camp in a single-storey wooden shack in the Salk Institute's car park. Armed with the latest high-performance liquid chromatography and gas-phase sequencing technologies, as well as thousands of sheep hypothalami, his team embarked on

Cell lines battle cancer

Large panels of human cancer cell lines, profiled at the DNA, RNA and chromosomal levels and tested for sensitivity to approved and potential drugs, will accelerate the search for new cancer therapies. [SEE ARTICLE P.570](#) & [LETTER P.603](#)

JOHN N. WEINSTEIN

The clinical trial is irreplaceable as a scientific testing ground for anticancer therapies. But sophisticated clinical trials are enormously expensive and difficult to perform, for logistical, regulatory, legal and ethical reasons. So, inevitably, we need model systems — in the laboratory and on the computer — to explore the molecular basis of drug activity and other aspects of cancer biology that cannot (and should not) be explored in people. Cultured cells are the most widely used of our model systems, despite their inability to reflect many aspects of a drug's behaviour in the human body. In this issue, Barretina *et al.*¹ (page 603) and Garnett *et al.*² (page 570) describe two large-scale resources that take cultured cancer cells and their pharmacology to the next level.

Barretina *et al.*¹ present what they term the Cancer Cell Line Encyclopedia (CCLE), an extensive compilation of gene expression, chromosome copy number and sequencing data on 947 publicly available human cancer cell lines of diverse lineages. They use two algorithms (naive Bayes and elastic net regression) to predict the sensitivity profiles of 479 of the cell lines to each of 24 different anticancer drugs, using cell lineage, genetic defects and gene expression as input variables. As a proof of principle for their approach, they present plausible correlations of drug activity with aberrations in particular genes (*IGF1R*, *AHR*, *NRAS* and *SLFN11*).

Garnett *et al.*² took a similar task, profiling a panel of several hundred diverse cancer cell lines for various genetic abnormalities, including point mutations, gene amplifications, gene deletions, microsatellite instability, frequently occurring DNA rearrangements and changes in gene expression. They then tested the sensitivity of those cells to 130 different anticancer agents (analysing between 275 and 507 cell lines per agent, for a total of 48,178 tests) and used the elastic net regression algorithm to predict drug sensitivity from the

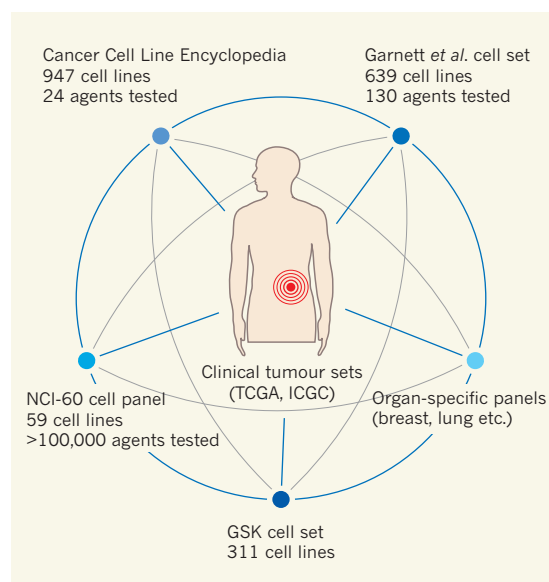


Figure 1 | The widening web of cancer cell lines. Barretina and colleagues' Cancer Cell Line Encyclopedia¹ and the cell panel presented by Garnett *et al.*² provide DNA, RNA and chromosome-level characterization, as well as drug-sensitivity profiles, of hundreds of cell lines derived from various types of cancer. The data sets can be cross-analysed with each other and with similar data from other panels of cultured cells such as the US National Cancer Institute's NCI-60 panel³, the GlaxoSmithKline (GSK) cancer-cell-line set⁷ and a number of organ-specific panels. The data can also be applied to generate and/or test hypotheses about clinical cancers, for example in projects of The Cancer Genome Atlas (TCGA)^{9,10} and the International Cancer Genome Consortium (ICGC)¹¹.

molecular profiles. Their analyses confirmed several expected gene–drug relationships and highlighted new associations, including a marked sensitivity of Ewing's sarcoma cells with the *EWS–FLI1* gene translocation to inhibitors of a group of DNA-repair proteins called PARPs.

Despite such reports of particular gene associations, and despite biologists' enduring preoccupation with strictly hypothesis-driven research, the principal significance of the two studies lies in the generic, largely hypothesis-independent data resources they provide to the research community³. Both data sets will be made publicly available and will undoubtedly be used by numerous investigators to generate or test their own hypotheses about particular

genes, proteins, pathways, cell lineages or drugs. The value of the data is increased by the fact that 496 cell lines are represented in both panels. The cell-culture conditions, methods used for molecular profiling and pharmacological assays differed somewhat between the two studies, but robust observations will often be reflected in data from both. Therefore, the differences in methodology can be considered either a disadvantage or a test of robustness.

In a similar fashion, information in the two data sets can be integrated with molecular and pharmacological profiles obtained from other panels of cell lines (Fig. 1). The pioneering panel of that type was the NCI-60, a set of 60 (now 59) human cancer cell lines from nine different tissues introduced⁴ in 1990 by the US National Cancer Institute (NCI) in Bethesda, Maryland, to screen for new anticancer agents. The NCI-60 cells have been profiled at the DNA, RNA, protein and chromosomal levels using a large number of technologies⁵. The panel is limited in number of cell types, but it has been used to test more than 100,000 chemically defined potential drugs and a larger number of natural-product extracts. Usefully, 55 of the 59 NCI-60 cell types are represented in at least one of the two panels reported in this issue.

Other publicly available information resources with which the data can be integrated include molecular and pharmacological data sets for breast⁶ and lung⁷ cancers, and for a panel of 311 cell lines⁸ that overlaps with those of Barretina *et al.*¹ and Garnett *et al.*².

All of the cell-line information resources described above are based on molecular profiles of untreated cells. They therefore assess or predict intrinsic cell sensitivity and resistance to various drugs or potential drugs. Another resource, the Connectivity Map⁹, takes a different approach: the cell lines are profiled both before and after treatment to assess how their molecular profiles respond to perturbation by the various agents tested. The complementary nature of the resulting data adds another useful layer of information for combined analyses.

The overarching challenge lies in relating the cell panels to clinical tumours. The Cancer Genome Atlas^{10,11}, a joint project of the NCI and the US National Human Genome Research Institute in Bethesda, provides such an opportunity, as does the International Cancer Genome Consortium project¹². Together, these two large-scale enterprises aim to generate comprehensive molecular profiles for more than 50 types of cancer from patients over the next few years. When the profiles reveal a DNA-level defect, or a difference in gene or protein expression, researchers will be able to check whether any of the cell-line data presented by Barretina and colleagues¹ or Garnett and colleagues² show a similar aberration. If so, the pertinent lines can be experimented with in ways that human participants cannot. For example, genes in the cells can be mutated selectively, the cells can be injected into mice to generate tumours, or expression of particular genes can be knocked down by RNA interference. Conversely, particular molecular aberrations or patterns of aberrations that predict the sensitivity of a cell line to a drug can be searched for in clinical cancers, to suggest

possible avenues for therapy. In other words, both 'bedside-to-bench' and 'bench-to-bedside' strategies for using the data resources are possible and productive.

The limitations of cell lines in culture as models for human pharmacology are well known: the cells have been removed from their interactions with other cell types, from their native tissue architecture, from the influence of cytokines and other cell-signalling molecules, and from the effects of drug distribution and metabolism in the body. Thus, indices of sensitivity and resistance in culture may not reflect the factors that influence a drug's action *in vivo*. Furthermore, many anticancer agents have dose-limiting organ toxicities that are not represented in model systems such as cultured cancer cells. Even with those caveats in mind, however, the molecular profiles presented by Barretina *et al.* and Garnett *et al.* provide highly useful resources for the generation and testing of hypotheses related to the grand goal of personalizing cancer medicine. As statistician George Box once wrote¹³, "all models are wrong but some are useful". ■

John N. Weinstein is in the Department of Bioinformatics and Computational Biology and the Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA.
e-mail: jweinste@mdanderson.org

1. Barretina, J. *et al.* *Nature* **483**, 603–607 (2012).
2. Garnett, M. J. *et al.* *Nature* **483**, 570–575 (2012).
3. Weinstein, J. N. *Curr. Opin. Pharmacol.* **2**, 361–365 (2002).
4. Shoemaker, R. H. *et al.* *Prog. Clin. Biol. Res.* **276**, 265–286 (1988).
5. Weinstein, J. N. *et al.* *Science* **275**, 343–349 (1997).
6. Neve, R. M. *et al.* *Cancer Cell* **10**, 515–527 (2006).
7. Sos, M. L. *et al.* *J. Clin. Invest.* **119**, 1727–1740 (2009).
8. Greshock, J. *et al.* *Cancer Res.* **70**, 3677–3686 (2010).
9. Lamb, J. *et al.* *Science* **313**, 1929–1935 (2006).
10. The Cancer Genome Atlas Research Network *Nature* **455**, 1061–1068 (2008).
11. The Cancer Genome Atlas Research Network *Nature* **474**, 609–615 (2011).
12. The International Cancer Genome Consortium *Nature* **464**, 993–998 (2010).
13. Box, G. E. P. in *Robustness in Statistics: Proceedings of a Workshop* (eds Launer, R. L. & Wilkinson, G. N.) (Academic, 1979).

PHOTONICS

Terahertz collisions

Intense laser fields can rip electrons from an atom and slam them back into it. By using intense terahertz radiation, this idea can be extended to electrons paired with 'holes' in a semiconductor. SEE LETTER P.580

RUPERT HUBER

It comes so naturally to us. To get a handle on new objects, say two marbles, children bounce them off each other to see whether they scatter or break. Scientists hunting for new phenomena in quantum physics are no less captivated by the idea of colliding atoms or elementary particles. Intense lasers have become a strong ally in that pursuit, ever since it was discovered that their oscillating electric field can rip an electron from an atom through quantum-mechanical tunnelling. In this process, the electron is rapidly accelerated by the laser's field until the field flips sign, forcing the electron to return to its starting point and collide with its parent ion. On this 'recollision', energetic photons are released, a phenomenon known as high-order-harmonic generation¹, and these photons carry key information about the electron and the atom. On page 580 of this issue, Zaks *et al.*² transfer this concept to a fascinating new class of low-energy recollision.

By using the powerful terahertz (10^{12} hertz) field of a free-electron laser, in place of the conventional near-infrared laser field that is

used to ionize atoms, the authors² managed to make the elementary components of an exciton³ bounce off each other in a semiconductor. The exciton is a particularly intriguing breed of quantum object that can emerge when a photon that has been absorbed by a semiconductor promotes an electron from the material's filled valence energy band into the empty conduction band, leaving a positively charged 'hole' behind. At low temperatures (usually below 10 kelvin), Coulomb attraction makes the electron swirl around the hole, just like an electron travels around the proton in a hydrogen atom. However, the binding energy of electrons in excitons is much lower than that of electrons in hydrogen atoms, and typically falls in the millielectronvolt range. Thus, field-induced ionization may be expected⁴ to occur at low frequencies, in the elusive terahertz window of the electromagnetic spectrum, which is located between the microwave and infrared regimes.

In their study, Zaks *et al.*² generated excitons in thin semiconductor layers using near-infrared laser light. To induce recollisions between the electrons and holes forming the excitons, they then exposed the semiconductor

to intense terahertz light that had amplitudes of 11.5 kilovolts per centimetre. The photon energy, $\hbar\omega_{\text{THz}}$ (where ω_{THz} is the terahertz angular frequency and \hbar is the reduced Planck constant) was tuned to a value well below the binding energy of the excitons (10 meV), to mimic high-order-harmonic generation in an atomic gas. But instead of detecting radiation generated at integer multiples (high-order harmonics) of the frequency of the driving terahertz field, the authors found a complementary signature of electron-hole recollisions: high-order-sideband generation⁵.

In this process, new frequencies, or sidebands, of radiation are generated and transmitted through the semiconductor material along with the near-infrared light that was used to create the excitons (Fig. 1). The photon energies of the sidebands ($\hbar\omega_{\text{sideband}}$) depend on the frequency of the driving terahertz field, as well as on that of the near-infrared light (ω_{NIR}): $\hbar\omega_{\text{sideband}} = \hbar\omega_{\text{NIR}} + n\hbar\omega_{\text{THz}}$, where n is the order of the sideband and is an even integer^{5–7}. Zaks *et al.* detected an impressive number of spectral lines, up to the eighteenth order. Most remarkably, they found that the intensity of the sidebands decays only slowly with increasing order. This slow decay is similar to that seen in high-order-harmonic generation, in which it is caused by an interference between the different trajectories of the electrons that are ripped from the atoms and accelerated¹.

To demonstrate that the detected sidebands originate from recollisions, Zaks and colleagues changed the polarization state of the terahertz field from linear to circular; light is linearly polarized if its electric field vibrates in a plane perpendicular to its direction of

The overarching challenge lies in relating the cell panels to clinical tumours. The Cancer Genome Atlas^{10,11}, a joint project of the NCI and the US National Human Genome Research Institute in Bethesda, provides such an opportunity, as does the International Cancer Genome Consortium project¹². Together, these two large-scale enterprises aim to generate comprehensive molecular profiles for more than 50 types of cancer from patients over the next few years. When the profiles reveal a DNA-level defect, or a difference in gene or protein expression, researchers will be able to check whether any of the cell-line data presented by Barretina and colleagues¹ or Garnett and colleagues² show a similar aberration. If so, the pertinent lines can be experimented with in ways that human participants cannot. For example, genes in the cells can be mutated selectively, the cells can be injected into mice to generate tumours, or expression of particular genes can be knocked down by RNA interference. Conversely, particular molecular aberrations or patterns of aberrations that predict the sensitivity of a cell line to a drug can be searched for in clinical cancers, to suggest

possible avenues for therapy. In other words, both 'bedside-to-bench' and 'bench-to-bedside' strategies for using the data resources are possible and productive.

The limitations of cell lines in culture as models for human pharmacology are well known: the cells have been removed from their interactions with other cell types, from their native tissue architecture, from the influence of cytokines and other cell-signalling molecules, and from the effects of drug distribution and metabolism in the body. Thus, indices of sensitivity and resistance in culture may not reflect the factors that influence a drug's action *in vivo*. Furthermore, many anticancer agents have dose-limiting organ toxicities that are not represented in model systems such as cultured cancer cells. Even with those caveats in mind, however, the molecular profiles presented by Barretina *et al.* and Garnett *et al.* provide highly useful resources for the generation and testing of hypotheses related to the grand goal of personalizing cancer medicine. As statistician George Box once wrote¹³, "all models are wrong but some are useful". ■

John N. Weinstein is in the Department of Bioinformatics and Computational Biology and the Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA.
e-mail: jweinste@mdanderson.org

1. Barretina, J. *et al.* *Nature* **483**, 603–607 (2012).
2. Garnett, M. J. *et al.* *Nature* **483**, 570–575 (2012).
3. Weinstein, J. N. *Curr. Opin. Pharmacol.* **2**, 361–365 (2002).
4. Shoemaker, R. H. *et al.* *Prog. Clin. Biol. Res.* **276**, 265–286 (1988).
5. Weinstein, J. N. *et al.* *Science* **275**, 343–349 (1997).
6. Neve, R. M. *et al.* *Cancer Cell* **10**, 515–527 (2006).
7. Sos, M. L. *et al.* *J. Clin. Invest.* **119**, 1727–1740 (2009).
8. Greshock, J. *et al.* *Cancer Res.* **70**, 3677–3686 (2010).
9. Lamb, J. *et al.* *Science* **313**, 1929–1935 (2006).
10. The Cancer Genome Atlas Research Network *Nature* **455**, 1061–1068 (2008).
11. The Cancer Genome Atlas Research Network *Nature* **474**, 609–615 (2011).
12. The International Cancer Genome Consortium *Nature* **464**, 993–998 (2010).
13. Box, G. E. P. in *Robustness in Statistics: Proceedings of a Workshop* (eds Launer, R. L. & Wilkinson, G. N.) (Academic, 1979).

PHOTONICS

Terahertz collisions

Intense laser fields can rip electrons from an atom and slam them back into it. By using intense terahertz radiation, this idea can be extended to electrons paired with 'holes' in a semiconductor. SEE LETTER P.580

RUPERT HUBER

It comes so naturally to us. To get a handle on new objects, say two marbles, children bounce them off each other to see whether they scatter or break. Scientists hunting for new phenomena in quantum physics are no less captivated by the idea of colliding atoms or elementary particles. Intense lasers have become a strong ally in that pursuit, ever since it was discovered that their oscillating electric field can rip an electron from an atom through quantum-mechanical tunnelling. In this process, the electron is rapidly accelerated by the laser's field until the field flips sign, forcing the electron to return to its starting point and collide with its parent ion. On this 'recollision', energetic photons are released, a phenomenon known as high-order-harmonic generation¹, and these photons carry key information about the electron and the atom. On page 580 of this issue, Zaks *et al.*² transfer this concept to a fascinating new class of low-energy recollision.

By using the powerful terahertz (10^{12} hertz) field of a free-electron laser, in place of the conventional near-infrared laser field that is

used to ionize atoms, the authors² managed to make the elementary components of an exciton³ bounce off each other in a semiconductor. The exciton is a particularly intriguing breed of quantum object that can emerge when a photon that has been absorbed by a semiconductor promotes an electron from the material's filled valence energy band into the empty conduction band, leaving a positively charged 'hole' behind. At low temperatures (usually below 10 kelvin), Coulomb attraction makes the electron swirl around the hole, just like an electron travels around the proton in a hydrogen atom. However, the binding energy of electrons in excitons is much lower than that of electrons in hydrogen atoms, and typically falls in the millielectronvolt range. Thus, field-induced ionization may be expected⁴ to occur at low frequencies, in the elusive terahertz window of the electromagnetic spectrum, which is located between the microwave and infrared regimes.

In their study, Zaks *et al.*² generated excitons in thin semiconductor layers using near-infrared laser light. To induce recollisions between the electrons and holes forming the excitons, they then exposed the semiconductor

to intense terahertz light that had amplitudes of 11.5 kilovolts per centimetre. The photon energy, $\hbar\omega_{\text{THz}}$ (where ω_{THz} is the terahertz angular frequency and \hbar is the reduced Planck constant) was tuned to a value well below the binding energy of the excitons (10 meV), to mimic high-order-harmonic generation in an atomic gas. But instead of detecting radiation generated at integer multiples (high-order harmonics) of the frequency of the driving terahertz field, the authors found a complementary signature of electron-hole recollisions: high-order-sideband generation⁵.

In this process, new frequencies, or sidebands, of radiation are generated and transmitted through the semiconductor material along with the near-infrared light that was used to create the excitons (Fig. 1). The photon energies of the sidebands ($\hbar\omega_{\text{sideband}}$) depend on the frequency of the driving terahertz field, as well as on that of the near-infrared light (ω_{NIR}): $\hbar\omega_{\text{sideband}} = \hbar\omega_{\text{NIR}} + n\hbar\omega_{\text{THz}}$, where n is the order of the sideband and is an even integer^{5–7}. Zaks *et al.* detected an impressive number of spectral lines, up to the eighteenth order. Most remarkably, they found that the intensity of the sidebands decays only slowly with increasing order. This slow decay is similar to that seen in high-order-harmonic generation, in which it is caused by an interference between the different trajectories of the electrons that are ripped from the atoms and accelerated¹.

To demonstrate that the detected sidebands originate from recollisions, Zaks and colleagues changed the polarization state of the terahertz field from linear to circular; light is linearly polarized if its electric field vibrates in a plane perpendicular to its direction of

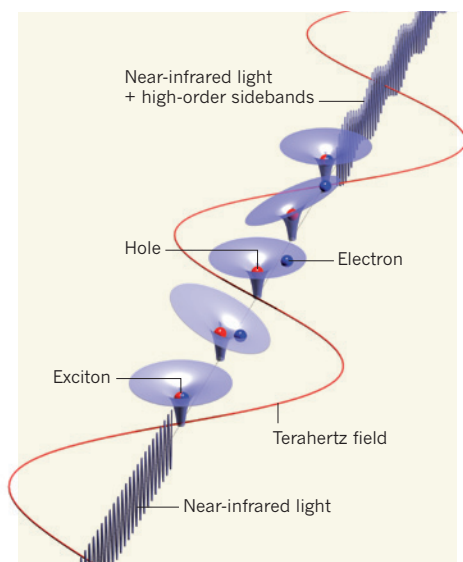


Figure 1 | Principle of high-order-sideband generation. Zaks *et al.*² created excitons — atom-like bound states of electrons and ‘holes’ — in a semiconductor (not shown) by shining near-infrared light on it. An intense terahertz electric field was used to skew the energy troughs (funnels) formed by electrostatic electron–hole attraction, pulling the electrons off the bound states and then back into them. Repeated ‘recollisions’ of the electrons with the holes caused new frequencies, or sidebands, of radiation to be produced and transmitted through the semiconductor, along with the near-infrared light.

propagation, and is circularly polarized if the field describes a helix about the propagation direction. The authors found that the sidebands were strongest for linear polarization and vanished for circular polarization, as expected from a classical analogy with the behaviour of atoms. In the case of linear polarization, the electron is forced into a large-amplitude oscillation in which it repeatedly collides with the hole. By contrast, with circular polarization, the electron is pulled by the field in constantly changing directions, never returning to the location of the hole from which it started its motion.

Under the conditions of the present experiment, the fairly robust analogy between excitons and atoms is remarkable. After all, excitons are transient elementary excitations, rather than conventional particles. As strange as it may sound, Zaks *et al.* actually do make excitations recollide. Furthermore, because solids typically consist of 10^{23} electrons and ions per cubic centimetre, which all mutually interact^{3,8}, it is clear that excitons may be influenced by one another, as well as by other elementary excitations. The authors acknowledge that such interactions could have affected their results. Scattering of the fastest electrons off neighbouring excitons, as well as vibrations in the semiconductor lattice structure, may have limited the number of high-order sidebands that could be resolved.

The experimental observation of electron–hole recollisions is great news for several reasons. First, high-order sidebands imprint a fast modulation on the near-infrared light. If the authors’ concept can be combined with state-of-the-art transistor amplifiers (electronic components that switch and amplify signals), acting as intense on-chip terahertz sources, it may be possible to drive ultrafast modulators at high rates, in the order of several terabits per second, for use in optical communication systems.

Second, and perhaps more importantly, the possibility of extending the physics of recollision beyond that of atoms and molecules fires the imagination, and should prompt the investigation of excitations in other materials. Third, the latest intense terahertz sources have reached amplitudes far above 1 megavolt per centimetre, concentrated in ultrashort flashes^{9,10}. It will be interesting to see how excitons behave under such extreme conditions. In particular, ultrafast electro-optic detectors may be able to directly monitor high-harmonic radiation with a temporal resolution faster than a single oscillation of the radiation field.

CANCER

Clinical trials unite mice and humans

Anticancer ‘co-clinical’ trials, in which mice carrying known mutations are treated in parallel with patients enrolled in a simultaneous clinical study, could help to improve therapeutic outcome. [SEE LETTER P.613](#)

LEISA JOHNSON

Advanced non-small-cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide. Over the past decade, treatment of patients with this disease has changed considerably to incorporate a more personalized approach. Many patients are tested for a key oncogenic mutation in their tumours’ DNA, and those who carry it are treated with a targeted therapy. Unfortunately, such tumours often display markedly diverse responses to the treatment. On page 613 of this issue, Chen and colleagues¹ use genetically engineered mouse models of NSCLC to interrogate the molecular complexity underlying the mixed therapeutic response. Moreover, they propose using these animal models to improve clinical success rates by conducting ‘co-clinical’ trials — mouse studies that mirror an ongoing clinical trial in patients whose tumours harbour the same driver mutations.

Certain mutations in the human *KRAS* gene

Will such detectors allow slow-motion ‘movies’ of collisions to be taken? Zaks and colleagues have set the stage for these and other questions of quantum many-body and solid-state physics to be addressed. The race is on. ■

Rupert Huber is in the Department of Physics, University of Regensburg, 93053 Regensburg, Germany.
e-mail: rupert.huber@physik.uni-regensburg.de

1. Corkum, P. B. & Krausz, F. *Nature Phys.* **3**, 381–387 (2007).
2. Zaks, B., Liu, R. B. & Sherwin, M. S. *Nature* **483**, 580–583 (2012).
3. Kira, M., Khitrova, G. & Gibbs, H. M. *Nature Mater.* **5**, 523–531 (2006).
4. Golde, D., Meier, T. & Koch, S. W. *Phys. Rev. B* **77**, 075330 (2008).
5. Yan, J.-Y. *Phys. Rev. B* **78**, 075204 (2008).
6. Kono, J. *et al. Phys. Rev. Lett.* **79**, 1758–1761 (1997).
7. Wagner, M. *et al. Appl. Phys. Lett.* **94**, 241105 (2009).
8. Axt, V. M. & Kuhn, T. *Rep. Prog. Phys.* **67**, 433–512 (2004).
9. Hebling, J., Yeh, K.-L., Hoffmann, M. C., Bartal, B. & Nelson, K. A. *J. Opt. Soc. Am. B* **25**, B6–B19 (2008).
10. Sell, A. *et al. Opt. Lett.* **33**, 2767–2769 (2008).

lead to an abnormal activation of the *KRAS* protein, which regulates several cellular signal-transduction pathways. These mutations occur in 20–30% of NSCLC cases and are predictive of a poor outcome in response to traditional anticancer drugs^{2,3}. There are no known inhibitors of oncogenic *KRAS*, but preclinical studies^{4,5} examining the inhibition of proteins such as MEK and PI3K — which act downstream of *KRAS* in the cellular pathways — have shown promise against *KRAS*-mutant NSCLC, particularly in combination with standard anticancer agents or with each other. These results have led to the initiation of several early-stage clinical trials interrogating new drug candidates against these pathways.

One ongoing clinical trial is designed to compare the activity of a standard cytotoxic drug (docetaxel) to that of the same agent in combination with a MEK inhibitor (selumetinib) in NSCLC patients whose tumours harbour *KRAS* mutations. Chen *et al.*¹ used this trial to explore whether selumetinib increases the efficacy of docetaxel in *Kras*-mutant

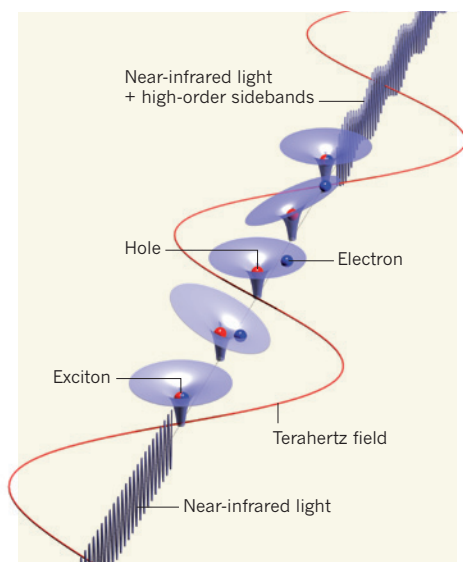


Figure 1 | Principle of high-order-sideband generation. Zaks *et al.*² created excitons — atom-like bound states of electrons and ‘holes’ — in a semiconductor (not shown) by shining near-infrared light on it. An intense terahertz electric field was used to skew the energy troughs (funnels) formed by electrostatic electron–hole attraction, pulling the electrons off the bound states and then back into them. Repeated ‘recollisions’ of the electrons with the holes caused new frequencies, or sidebands, of radiation to be produced and transmitted through the semiconductor, along with the near-infrared light.

propagation, and is circularly polarized if the field describes a helix about the propagation direction. The authors found that the sidebands were strongest for linear polarization and vanished for circular polarization, as expected from a classical analogy with the behaviour of atoms. In the case of linear polarization, the electron is forced into a large-amplitude oscillation in which it repeatedly collides with the hole. By contrast, with circular polarization, the electron is pulled by the field in constantly changing directions, never returning to the location of the hole from which it started its motion.

Under the conditions of the present experiment, the fairly robust analogy between excitons and atoms is remarkable. After all, excitons are transient elementary excitations, rather than conventional particles. As strange as it may sound, Zaks *et al.* actually do make excitations recollide. Furthermore, because solids typically consist of 10^{23} electrons and ions per cubic centimetre, which all mutually interact^{3,8}, it is clear that excitons may be influenced by one another, as well as by other elementary excitations. The authors acknowledge that such interactions could have affected their results. Scattering of the fastest electrons off neighbouring excitons, as well as vibrations in the semiconductor lattice structure, may have limited the number of high-order sidebands that could be resolved.

The experimental observation of electron–hole recollisions is great news for several reasons. First, high-order sidebands imprint a fast modulation on the near-infrared light. If the authors’ concept can be combined with state-of-the-art transistor amplifiers (electronic components that switch and amplify signals), acting as intense on-chip terahertz sources, it may be possible to drive ultrafast modulators at high rates, in the order of several terabits per second, for use in optical communication systems.

Second, and perhaps more importantly, the possibility of extending the physics of recollision beyond that of atoms and molecules fires the imagination, and should prompt the investigation of excitations in other materials. Third, the latest intense terahertz sources have reached amplitudes far above 1 megavolt per centimetre, concentrated in ultrashort flashes^{9,10}. It will be interesting to see how excitons behave under such extreme conditions. In particular, ultrafast electro-optic detectors may be able to directly monitor high-harmonic radiation with a temporal resolution faster than a single oscillation of the radiation field.

CANCER

Clinical trials unite mice and humans

Anticancer ‘co-clinical’ trials, in which mice carrying known mutations are treated in parallel with patients enrolled in a simultaneous clinical study, could help to improve therapeutic outcome. [SEE LETTER P.613](#)

LEISA JOHNSON

Advanced non-small-cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide. Over the past decade, treatment of patients with this disease has changed considerably to incorporate a more personalized approach. Many patients are tested for a key oncogenic mutation in their tumours’ DNA, and those who carry it are treated with a targeted therapy. Unfortunately, such tumours often display markedly diverse responses to the treatment. On page 613 of this issue, Chen and colleagues¹ use genetically engineered mouse models of NSCLC to interrogate the molecular complexity underlying the mixed therapeutic response. Moreover, they propose using these animal models to improve clinical success rates by conducting ‘co-clinical’ trials — mouse studies that mirror an ongoing clinical trial in patients whose tumours harbour the same driver mutations.

Certain mutations in the human *KRAS* gene

Will such detectors allow slow-motion ‘movies’ of collisions to be taken? Zaks and colleagues have set the stage for these and other questions of quantum many-body and solid-state physics to be addressed. The race is on. ■

Rupert Huber is in the Department of Physics, University of Regensburg, 93053 Regensburg, Germany.
e-mail: rupert.huber@physik.uni-regensburg.de

1. Corkum, P. B. & Krausz, F. *Nature Phys.* **3**, 381–387 (2007).
2. Zaks, B., Liu, R. B. & Sherwin, M. S. *Nature* **483**, 580–583 (2012).
3. Kira, M., Khitrova, G. & Gibbs, H. M. *Nature Mater.* **5**, 523–531 (2006).
4. Golde, D., Meier, T. & Koch, S. W. *Phys. Rev. B* **77**, 075330 (2008).
5. Yan, J.-Y. *Phys. Rev. B* **78**, 075204 (2008).
6. Kono, J. *et al. Phys. Rev. Lett.* **79**, 1758–1761 (1997).
7. Wagner, M. *et al. Appl. Phys. Lett.* **94**, 241105 (2009).
8. Axt, V. M. & Kuhn, T. *Rep. Prog. Phys.* **67**, 433–512 (2004).
9. Hebling, J., Yeh, K.-L., Hoffmann, M. C., Bartal, B. & Nelson, K. A. *J. Opt. Soc. Am. B* **25**, B6–B19 (2008).
10. Sell, A. *et al. Opt. Lett.* **33**, 2767–2769 (2008).

lead to an abnormal activation of the *KRAS* protein, which regulates several cellular signal-transduction pathways. These mutations occur in 20–30% of NSCLC cases and are predictive of a poor outcome in response to traditional anticancer drugs^{2,3}. There are no known inhibitors of oncogenic *KRAS*, but preclinical studies^{4,5} examining the inhibition of proteins such as MEK and PI3K — which act downstream of *KRAS* in the cellular pathways — have shown promise against *KRAS*-mutant NSCLC, particularly in combination with standard anticancer agents or with each other. These results have led to the initiation of several early-stage clinical trials interrogating new drug candidates against these pathways.

One ongoing clinical trial is designed to compare the activity of a standard cytotoxic drug (docetaxel) to that of the same agent in combination with a MEK inhibitor (selumetinib) in NSCLC patients whose tumours harbour *KRAS* mutations. Chen *et al.*¹ used this trial to explore whether selumetinib increases the efficacy of docetaxel in *Kras*-mutant

tumours in mice, and how the presence of additional genetic alterations — frequently found in close association with *KRAS* mutations in humans — may influence therapeutic response. Specifically, the authors used three genetically engineered mouse models of NSCLC to examine the role of mutations in the tumour-suppressor genes *p53* and *Lkb1*. In these mice, the expression of *Kras*, of *Kras* and *p53*, or of *Kras* and *Lkb1* could be precisely manipulated so that the animals developed multifocal disease that closely emulated human NSCLC, with each lesion progressing at an independent rate, as previous studies have shown^{6–8}.

Using magnetic resonance imaging and microscopy to assess tumour-cell proliferation and death, Chen and colleagues¹ found that mutation of either *p53* or *Lkb1* in *Kras*-mutant tumours in mice significantly diminished the initial effect of docetaxel on the tumours. The addition of selumetinib enhanced docetaxel's effect on *Kras*- and *Kras/p53*-mutant tumours, and improved progression-free survival — the time elapsed between treatment initiation and tumour progression or death from any cause — in both mouse models. By contrast, *Kras/Lkb1*-mutant tumours were inherently resistant to this combination therapy. Therefore, the docetaxel–selumetinib combination may be less effective in patients with tumours carrying mutations in both *KRAS* and *LKB1*.

To measure metabolic changes in tumours as a possible surrogate for defining early response to the therapy, the authors¹ injected a radiolabelled glucose analogue (¹⁸F-fluoro-2-deoxyglucose, FDG) into the mice and traced its concentration in the tumours with positron emission tomography (PET), a powerful imaging technique. They found that *Kras/p53*- and *Kras/Lkb1*-mutant tumours have an overall higher FDG uptake than *Kras*-only mutant tumours. Chen *et al.*¹ note that one partial explanation for this may be increased expression of GLUT1 — a protein that controls glucose uptake into cells — in the *Kras/Lkb1* mutant tumours. Alternatively, the differences in FDG uptake may also reflect disease subtype, stage or both. Importantly, the authors translated these observations to humans by finding a significant correlation between LKB1 expression and FDG avidity in human NSCLC.

The researchers then explored the usefulness of FDG–PET to determine tumour metabolic changes following short-term therapeutic intervention in mice, and found that the combination of docetaxel and selumetinib reduced tumour metabolic activity only in the *Kras*- and *Kras/p53*-mutant mice.

These results agree with the authors' microscopic study of tumour-cell proliferation and death rates, and suggest that serial

FDG–PET imaging may be useful clinically in predicting antitumour efficacy and patient outcome in *KRAS*-mutant NSCLC treated with this combination therapy.

Of note is that most, if not all, of the lesions examined by Chen *et al.*¹ in the *Kras*- and *Kras/p53*-mutant mice represent an earlier stage of the disease than that typically evaluated in initial, exploratory NSCLC clinical trials for therapeutics. Earlier-stage disease is also generally more responsive to therapy, with better long-term outcomes. So, serial FDG–PET in this context could be more useful if future studies could directly correlate FDG uptake with microscopic analyses detailing tumour stage, subtype, genotype and therapeutic response for each lesion.

Previous reports^{9,10} have suggested that elevated FDG uptake in human lung tumours predicts poor outcome in response to conventional anticancer drugs. The results of Chen and colleagues' study¹ suggest that this may also extend to targeted therapies, specifically to the treatment of *KRAS/LKB1*-mutant NSCLC with a combination of selumetinib and docetaxel. The authors propose that FDG–PET imaging may be used to identify patients who are more likely to respond and, therefore, to have a better long-term outcome. However, such an approach should be used in combination with other strategies to facilitate patient stratification, as FDG avidity was insufficient at predicting response in all *Kras/Lkb1*-mutant mice in the authors' report.

Overall, Chen and colleagues' work highlights the vital need to develop improved preclinical and clinical tools to follow and characterize individual tumours throughout the course of treatment. Several tumour attributes should be simultaneously correlated with therapeutic response over time to better understand resistance mechanisms. To this end, we need to develop more sophisticated reporter molecules and *in vivo* imaging modalities than FDG–PET to interrogate key cellular pathways in a dynamic way, and in real time.

The high failure rate of clinical trials for the treatment of late-stage diseases — particularly cancer¹¹ — underscores the need for improved preclinical models, as well as their translation into clinical-trial design, analysis and predictions. Chen and colleagues¹ present a compelling case for conducting co-clinical trials in genetically engineered mice or in other well-validated, relevant model systems such as patient-derived xenografts (in which a piece of the patient's tumour, or cells derived from it, are transplanted into a laboratory mouse). If done properly, co-clinical trials may help to identify predictive genetic markers that can be validated in real time using samples from patients enrolled in a concurrent clinical trial. These integrated data sets may ultimately be better at predicting the results of the concurrent clinical studies, as well as providing, on the basis of the cancer's genetic profile, a rationale



50 Years Ago

Applications are invited for a scholarship sponsored by the Worshipful Company of Gardeners, and open to young gardeners who are undergoing or have completed training at the Royal Horticultural Society's Gardens, Wisley, or elsewhere, and who will have had at least four years' practical experience in horticulture ... The scholarship is restricted to male candidates who are unmarried and undertake to remain so during the tenure of the scholarship. The scholarship will be tenable for two years, beginning October 1, and is valued at £300 per annum.

From *Nature* 31 March 1962

100 Years Ago

The metals occurring most frequently are gold and copper. The former is much more widely distributed than the latter, and must have been the first metal to be known in many regions. It is, however, one of the most worthless metals for practical purposes, so that until the rise of Greek and Roman civilisation but little use was made of it. Copper, too, we only find in use to a very limited extent, as it was not well suited for the construction of weapons or useful implements. On the other hand, its alloy with tin afforded a metal which in many physical properties could only be surpassed by iron or steel. According to the views of several ancient writers, Lucretius and Poseidonius, so momentous a discovery as that of metals contained in ores must needs have been brought about by no uncommon cause. According to them, a conflagration consumed forests which covered the outcrop of metalliferous veins, reducing the metals and bringing them to the notice of man, but there are no grounds for such inference.

From *Nature* 28 March 1912

➔ **NATURE.COM**
For more on
co-clinical
trials, see:
go.nature.com/ayaolt

tumours in mice, and how the presence of additional genetic alterations — frequently found in close association with *KRAS* mutations in humans — may influence therapeutic response. Specifically, the authors used three genetically engineered mouse models of NSCLC to examine the role of mutations in the tumour-suppressor genes *p53* and *Lkb1*. In these mice, the expression of *Kras*, of *Kras* and *p53*, or of *Kras* and *Lkb1* could be precisely manipulated so that the animals developed multifocal disease that closely emulated human NSCLC, with each lesion progressing at an independent rate, as previous studies have shown^{6–8}.

Using magnetic resonance imaging and microscopy to assess tumour-cell proliferation and death, Chen and colleagues¹ found that mutation of either *p53* or *Lkb1* in *Kras*-mutant tumours in mice significantly diminished the initial effect of docetaxel on the tumours. The addition of selumetinib enhanced docetaxel's effect on *Kras*- and *Kras/p53*-mutant tumours, and improved progression-free survival — the time elapsed between treatment initiation and tumour progression or death from any cause — in both mouse models. By contrast, *Kras/Lkb1*-mutant tumours were inherently resistant to this combination therapy. Therefore, the docetaxel–selumetinib combination may be less effective in patients with tumours carrying mutations in both *KRAS* and *LKB1*.

To measure metabolic changes in tumours as a possible surrogate for defining early response to the therapy, the authors¹ injected a radiolabelled glucose analogue (¹⁸F-fluoro-2-deoxyglucose, FDG) into the mice and traced its concentration in the tumours with positron emission tomography (PET), a powerful imaging technique. They found that *Kras/p53*- and *Kras/Lkb1*-mutant tumours have an overall higher FDG uptake than *Kras*-only mutant tumours. Chen *et al.*¹ note that one partial explanation for this may be increased expression of GLUT1 — a protein that controls glucose uptake into cells — in the *Kras/Lkb1* mutant tumours. Alternatively, the differences in FDG uptake may also reflect disease subtype, stage or both. Importantly, the authors translated these observations to humans by finding a significant correlation between LKB1 expression and FDG avidity in human NSCLC.

The researchers then explored the usefulness of FDG–PET to determine tumour metabolic changes following short-term therapeutic intervention in mice, and found that the combination of docetaxel and selumetinib reduced tumour metabolic activity only in the *Kras*- and *Kras/p53*-mutant mice.

These results agree with the authors' microscopic study of tumour-cell proliferation and death rates, and suggest that serial

FDG–PET imaging may be useful clinically in predicting antitumour efficacy and patient outcome in *KRAS*-mutant NSCLC treated with this combination therapy.

Of note is that most, if not all, of the lesions examined by Chen *et al.*¹ in the *Kras*- and *Kras/p53*-mutant mice represent an earlier stage of the disease than that typically evaluated in initial, exploratory NSCLC clinical trials for therapeutics. Earlier-stage disease is also generally more responsive to therapy, with better long-term outcomes. So, serial FDG–PET in this context could be more useful if future studies could directly correlate FDG uptake with microscopic analyses detailing tumour stage, subtype, genotype and therapeutic response for each lesion.

Previous reports^{9,10} have suggested that elevated FDG uptake in human lung tumours predicts poor outcome in response to conventional anticancer drugs. The results of Chen and colleagues' study¹ suggest that this may also extend to targeted therapies, specifically to the treatment of *KRAS/LKB1*-mutant NSCLC with a combination of selumetinib and docetaxel. The authors propose that FDG–PET imaging may be used to identify patients who are more likely to respond and, therefore, to have a better long-term outcome. However, such an approach should be used in combination with other strategies to facilitate patient stratification, as FDG avidity was insufficient at predicting response in all *Kras/Lkb1*-mutant mice in the authors' report.

Overall, Chen and colleagues' work highlights the vital need to develop improved preclinical and clinical tools to follow and characterize individual tumours throughout the course of treatment. Several tumour attributes should be simultaneously correlated with therapeutic response over time to better understand resistance mechanisms. To this end, we need to develop more sophisticated reporter molecules and *in vivo* imaging modalities than FDG–PET to interrogate key cellular pathways in a dynamic way, and in real time.

The high failure rate of clinical trials for the treatment of late-stage diseases — particularly cancer¹¹ — underscores the need for improved preclinical models, as well as their translation into clinical-trial design, analysis and predictions. Chen and colleagues¹ present a compelling case for conducting co-clinical trials in genetically engineered mice or in other well-validated, relevant model systems such as patient-derived xenografts (in which a piece of the patient's tumour, or cells derived from it, are transplanted into a laboratory mouse). If done properly, co-clinical trials may help to identify predictive genetic markers that can be validated in real time using samples from patients enrolled in a concurrent clinical trial. These integrated data sets may ultimately be better at predicting the results of the concurrent clinical studies, as well as providing, on the basis of the cancer's genetic profile, a rationale



50 Years Ago

Applications are invited for a scholarship sponsored by the Worshipful Company of Gardeners, and open to young gardeners who are undergoing or have completed training at the Royal Horticultural Society's Gardens, Wisley, or elsewhere, and who will have had at least four years' practical experience in horticulture ... The scholarship is restricted to male candidates who are unmarried and undertake to remain so during the tenure of the scholarship. The scholarship will be tenable for two years, beginning October 1, and is valued at £300 per annum.

From *Nature* 31 March 1962

100 Years Ago

The metals occurring most frequently are gold and copper. The former is much more widely distributed than the latter, and must have been the first metal to be known in many regions. It is, however, one of the most worthless metals for practical purposes, so that until the rise of Greek and Roman civilisation but little use was made of it. Copper, too, we only find in use to a very limited extent, as it was not well suited for the construction of weapons or useful implements. On the other hand, its alloy with tin afforded a metal which in many physical properties could only be surpassed by iron or steel. According to the views of several ancient writers, Lucretius and Poseidonius, so momentous a discovery as that of metals contained in ores must needs have been brought about by no uncommon cause. According to them, a conflagration consumed forests which covered the outcrop of metalliferous veins, reducing the metals and bringing them to the notice of man, but there are no grounds for such inference.

From *Nature* 28 March 1912

for the observed differences in therapeutic response. Moreover, such coordinated processes could serve to inform the analysis and design of both current and future clinical trials, with a goal of increasing clinical success rates and decreasing health-care costs. ■

Leisa Johnson is at Genentech, South San Francisco, California 94080, USA.

e-mail: johnson.leisa@gene.com

1. Chen, Z. *et al.* *Nature* **483**, 613–617 (2012).
2. Douillard, J. Y. *et al.* *J. Clin. Oncol.* **28**, 744–752 (2010).
3. Mascaux, C. *et al.* *Br. J. Cancer* **92**, 131–139 (2005).
4. Ji, H. *et al.* *Cancer Res.* **67**, 4933–4939 (2007).
5. Hoeflich, K. P. *et al.* *Cancer Res.* **72**, 210–219 (2012).
6. Jackson, E. L. *et al.* *Cancer Res.* **65**, 10280–10288 (2005).

7. Singh, M. *et al.* *Nature Biotechnol.* **28**, 585–593 (2010).
8. Ji, H. *et al.* *Nature* **448**, 807–810 (2007).
9. Berghmans, T. *et al.* *Thorac. Oncol.* **3**, 6–12 (2008).
10. Erasmus, J. J., Rohren, E. & Swisher, S. G. *Proc. Am. Thorac. Soc.* **6**, 171–179 (2009).
11. Kola, I. & Landis, J. *Nature Rev. Drug Discov.* **3**, 711–716 (2004).

The author declares competing financial interests. See go.nature.com/qhuvvg for details.

PALAEOCLIMATE

Tahitian record suggests Antarctic collapse

The exact origin, timing and amplitude of a rapid period of sea-level rise known as meltwater pulse 1A, about 14,500 years ago, have remained unclear. An analysis of coral samples from Tahiti delivers some answers. [SEE ARTICLE P.559](#)

ROBERT E. KOPP

Understanding ice-sheet dynamics in Earth's past is central to testing our understanding of how ice sheets might behave in the future, and therefore to improving projections of future sea-level rise. The most recent episode of great sea-level change occurred as a result of the melting of the massive ice sheets of the Last Glacial Maximum, which ended about 19,000 years ago¹. In 1989, a record of sea-level change at Barbados — based on fossil corals in sediment cores and the depth distribution of their modern descendants — provided a detailed perspective on the interval from 19,000 to 8,000 years ago, when sea level at Barbados rose from about 120 metres to roughly 20 m below its current level². This record revealed that the sea-level rise was not smooth, but was instead punctuated by a sharp increase — estimated to be about 24 m in less than 1,000 years — dubbed meltwater pulse 1A

(MWP-1A). Evidence for MWP-1A was subsequently found in records from a number of other sites, including Tahiti, Hawaii and the Sunda Shelf in southeast Asia. But, until now, no locality had produced a record that matched the bathymetric and chronological precision of the Barbados record.

On page 559 of this issue, Deschamps *et al.*³ present a new sea-level record of MWP-1A from Tahiti that ends this data drought. Their results are based on fossil corals and vermetid gastropods (a type of sea snail) collected during the Integrated Ocean Drilling Program Expedition 310, which cored 37 boreholes from sites off the island's northern, western and southern coasts⁴. This record allows a fresh look at some key questions about MWP-1A. When, precisely, did it take place? How big was it? Where did the meltwater come from? And what was the relationship between MWP-1A and climate changes taking place during the end of the ice age?

Using the radioactive decay of uranium into thorium in the fossil corals as a clock, Deschamps *et al.* find that MWP-1A began between 14,650 and 14,500 years ago, and ended by 14,310 years ago, after a jump in sea level of 12–22 m (probably 14–18 m). Thus, in Tahiti, MWP-1A produced a rate of sea-level rise of at least 3.5 m per century, probably about 5 m per century, and conceivably more than twice as fast. These rates are much greater than current rates of mean global sea-level rise — about 0.3 m per century — and significantly more than most high-end estimates of twenty-first-century rates of global sea-level rise, which are generally below 2 m per century.

These dates for sea-level rise in Tahiti are inconsistent with older Barbados reconstructions, such as those discussed by Deschamps *et al.*, that ascribe MWP-1A there to a sea-level jump and data gap between two cores at 14,100–13,600 years ago. They are, however, consistent with the most recent and data-rich reconstruction⁵ of sea level at Barbados during MWP-1A, presented at the 2010 International Conference on Paleoclimatology. This reconstruction suggests that MWP-1A began there around 14,500 to 14,700 years ago and had a magnitude of about 20 m (slightly larger than the roughly 15 m Deschamps *et al.* estimate when they re-interpret the older, sparser record in light of their dates).

The only mechanism that can generate metres per century of global sea-level rise is melting ice sheets. Since its discovery,

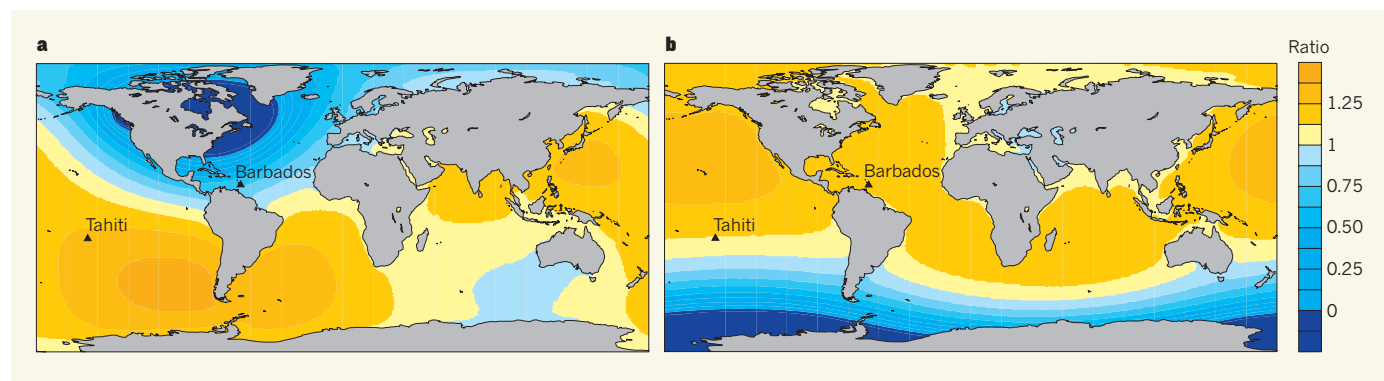


Figure 1 | Sea-level fingerprints of meltwater pulse 1A (MWP-1A). Predictions of the ratio of local sea-level rise to global sea-level rise associated with: **a**, melting of the southern one-third of the North America's Laurentide Ice Sheet; and **b**, melting of the West Antarctic Ice Sheet, both at the onset of

the MWP-1A event. The similarity between the sea-level record of MWP-1A from Tahiti obtained by Deschamps *et al.*³ and another record⁵ of the event from Barbados points to a largely Antarctic source for MWP-1A. (Figure modified from ref. 9.)

MWP-1A has commonly been viewed as a short-lived acceleration within the long-term decline of North America's Laurentide Ice Sheet². At the end of the Last Glacial Maximum, the Laurentide is estimated^{1,6} to have contained enough water to raise global sea level by about 70 m; by roughly 7,000 years ago it had almost all disappeared⁷. Although it is natural to interpret MWP-1A as a manifestation of that decline, several problems have led many researchers to challenge the hypothesis of a primarily Laurentide source, and suggest instead a major Antarctic contribution⁸.

With multiple well-dated records, it should be possible to 'fingerprint' the meltwater sources⁹ (Fig. 1). When an ice sheet melts, a sizeable amount of water is redistributed from a fairly concentrated source (the ice sheet) to a distributed one (the ocean). This mass redistribution reshapes Earth's gravitational field, lessens the flexure of the lithosphere (Earth's rigid outermost layer) in the vicinity of the ice sheet and alters the rate and orientation of Earth's rotation. The net effect is an initial sea-level fall near a melting ice sheet and enhanced sea-level rise far from the ice sheet. Thus, Laurentide melt would have caused about 40% less sea-level rise in Barbados than in Tahiti, whereas Antarctic melt would have caused similar amounts of sea-level rise at both localities⁹. The similarity of sea-level rise at Barbados and Tahiti is most consistent with a predominantly Antarctic source, and is difficult to reconcile with a purely Laurentide one.

Deschamps and colleagues' Tahiti chronology and the most recent Barbados chronology⁵ of MWP-1A indicate that the meltwater pulse started at around the same time as a period of warming in the Northern Hemisphere known as the Bølling, an episode of cooling in the Southern Hemisphere called the Antarctic cold reversal, and an associated strengthening of the Atlantic meridional overturning circulation (AMOC)¹⁰. Through the AMOC (the 'conveyor belt' that carries warm, upper Atlantic Ocean water to high northern latitudes and returns cold, deep waters to the south), Antarctic melt and the northern Bølling warming could have acted as feedbacks on one another. The introduction of fresh water into the Southern Ocean would have strengthened the AMOC, leading to an attendant northern warming and southern cooling¹¹. Conversely, a warmer Northern Hemisphere would have promoted Northern Hemisphere ice-sheet melting, causing a sea-level rise that would have destabilized marine-based parts of the Antarctic ice sheet.

The evidence from sea-level fingerprints for a primarily Antarctic source of MWP-1A is unlikely to be the last word. Although geochemical records are consistent with less than about 5 m of melt sourced from the Laurentide¹², geologists working in both East and West Antarctica have had difficulty finding evidence for an ice-sheet retreat of the

necessary scale and as early as required to explain MWP-1A^{13,14}. But for the moment, the geographical patterns seen in the sea-level records of MWP-1A argue that the event was caused predominantly by rapid Antarctic melting. This evidence for Antarctic instability emphasizes that, although a negative local sea-level feedback may reduce the instability of marine-based ice sheets¹⁵, this feedback cannot be regarded as a guarantee against the collapse of the marine-based sectors of the Antarctic ice-sheet in the face of a warmer and rising sea. There is enough marine-based ice remaining in the West Antarctic Ice Sheet today to raise global sea level by about 3.3 m (ref. 16). The example of MWP-1A serves as a reminder of the risk the ice sheet poses to the world's coasts. ■

Robert E. Kopp is in the Department of Earth and Planetary Sciences and Rutgers Energy Institute, Rutgers University, Piscataway, New Jersey 08854, USA.
e-mail: robert.kopp@rutgers.edu

HUMAN EVOLUTION

Those feet in ancient times

A fossil foot found in Ethiopia suggests that human ancestors that walked on two feet and also ably climbed trees existed until 3.4 million years ago, adding evidence for locomotor diversity during early human evolution. [SEE ARTICLE P.565](#)

DANIEL E. LIEBERMAN

The limitations of the fossil record leave ample room for debate about human origins. But most palaeoanthropologists agree that selection for bipedalism was instrumental in setting the human lineage on its separate evolutionary path from the chimpanzee lineage. And, as with any journey, it was probably sensible for our ancestors to put their best foot forward when starting out. The big question is, what kind of foot? On page 565 of this issue, Haile-Selassie and colleagues¹ present findings from a partial foot fossil which suggest that the feet of early hominins (species more closely related to humans than to chimpanzees), and hence their locomotor behaviour, were more diverse than was previously thought, and that the diversity lasted for much longer than was thought.

Human feet are remarkably different from those of apes² (Fig. 1). We have long, hefty big toes whose orientation does not diverge from that of the other toes, which are shorter and straighter than in other primates. Our feet also have a large, stable heel, for striking the ground

when we walk, and a well-developed arch that stiffens the middle of the foot and transfers the body's weight inward towards the base of the big toe, helping to push the body forward and upward at the end of stance.

Many of these distinctive features are also present in foot bones belonging to species of *Australopithecus*, a diverse genus of hominin that lived in Africa from about 4.4 million to 1.3 million years ago³. An absence of fossil feet older than those of *Australopithecus* led palaeoanthropologists to believe that human-like feet helped guide the way in human evolution, by enabling early hominins to walk effectively as bipeds, even while they retained some features that helped them to climb trees. In addition, the origin of the genus *Homo*, to which modern humans belong, was thought to have involved only minor modifications to foot anatomy, perhaps to improve our ancestors' ability to run long distances, although at the expense of climbing⁴.

Recent discoveries have made that simple narrative more complex. Most importantly, a fossil foot from *Ardipithecus ramidus*⁵, a 4.4-million-year-old species of hominin,

- Clark, P. U. *et al. Science* **325**, 710–714 (2009).
- Fairbanks, R. G. A. *Nature* **342**, 637–642 (1989).
- Deschamps, P. *et al. Nature* **483**, 559–564 (2012).
- <http://publications.iodp.org/proceedings/310/310title.htm>
- Mortlock, R. *et al. in 10th Int. Conf. Paleoceanography*, La Jolla (2010).
- Peltier, W. R. *Annu. Rev. Earth Planet. Sci.* **32**, 111–149 (2004).
- Carlson, A. E., Stoner, J. S., Donnelly, J. P. & Hillaire-Marcel, C. *Geology* **36**, 359–362 (2008).
- Clark, P. U. *et al. Paleoceanography* **11**, 563–577 (1996).
- Clark, P. U., Mitrovica, J. X., Milne, G. A. & Tamisiea, M. E. *Science* **295**, 2438–2441 (2002).
- Barker, S. *et al. Nature* **457**, 1097–1102 (2009).
- Weaver, A. J., Saenko, O. A., Clark, P. U. & Mitrovica, J. X. *Science* **299**, 1709–1713 (2003).
- Carlson, A. E. *Quat. Sci. Rev.* **28**, 1625–1630 (2009).
- Mackintosh, A. *et al. Nature Geosci.* **4**, 195–202 (2011).
- Bentley, M. J. *et al. Geology* **38**, 411–414 (2010).
- Gomez, N., Mitrovica, J. X., Huybers, P. & Clark, P. U. *Nature Geosci.* **3**, 850–853 (2010).
- Bamber, J. L., Riva, R. E. M., Vermeersen, B. L. A. & LeBrocq, A. M. *Science* **324**, 901–903 (2009).

MWP-1A has commonly been viewed as a short-lived acceleration within the long-term decline of North America's Laurentide Ice Sheet². At the end of the Last Glacial Maximum, the Laurentide is estimated^{1,6} to have contained enough water to raise global sea level by about 70 m; by roughly 7,000 years ago it had almost all disappeared⁷. Although it is natural to interpret MWP-1A as a manifestation of that decline, several problems have led many researchers to challenge the hypothesis of a primarily Laurentide source, and suggest instead a major Antarctic contribution⁸.

With multiple well-dated records, it should be possible to 'fingerprint' the meltwater sources⁹ (Fig. 1). When an ice sheet melts, a sizeable amount of water is redistributed from a fairly concentrated source (the ice sheet) to a distributed one (the ocean). This mass redistribution reshapes Earth's gravitational field, lessens the flexure of the lithosphere (Earth's rigid outermost layer) in the vicinity of the ice sheet and alters the rate and orientation of Earth's rotation. The net effect is an initial sea-level fall near a melting ice sheet and enhanced sea-level rise far from the ice sheet. Thus, Laurentide melt would have caused about 40% less sea-level rise in Barbados than in Tahiti, whereas Antarctic melt would have caused similar amounts of sea-level rise at both localities⁹. The similarity of sea-level rise at Barbados and Tahiti is most consistent with a predominantly Antarctic source, and is difficult to reconcile with a purely Laurentide one.

Deschamps and colleagues' Tahiti chronology and the most recent Barbados chronology⁵ of MWP-1A indicate that the meltwater pulse started at around the same time as a period of warming in the Northern Hemisphere known as the Bølling, an episode of cooling in the Southern Hemisphere called the Antarctic cold reversal, and an associated strengthening of the Atlantic meridional overturning circulation (AMOC)¹⁰. Through the AMOC (the 'conveyor belt' that carries warm, upper Atlantic Ocean water to high northern latitudes and returns cold, deep waters to the south), Antarctic melt and the northern Bølling warming could have acted as feedbacks on one another. The introduction of fresh water into the Southern Ocean would have strengthened the AMOC, leading to an attendant northern warming and southern cooling¹¹. Conversely, a warmer Northern Hemisphere would have promoted Northern Hemisphere ice-sheet melting, causing a sea-level rise that would have destabilized marine-based parts of the Antarctic ice sheet.

The evidence from sea-level fingerprints for a primarily Antarctic source of MWP-1A is unlikely to be the last word. Although geochemical records are consistent with less than about 5 m of melt sourced from the Laurentide¹², geologists working in both East and West Antarctica have had difficulty finding evidence for an ice-sheet retreat of the

necessary scale and as early as required to explain MWP-1A^{13,14}. But for the moment, the geographical patterns seen in the sea-level records of MWP-1A argue that the event was caused predominantly by rapid Antarctic melting. This evidence for Antarctic instability emphasizes that, although a negative local sea-level feedback may reduce the instability of marine-based ice sheets¹⁵, this feedback cannot be regarded as a guarantee against the collapse of the marine-based sectors of the Antarctic ice-sheet in the face of a warmer and rising sea. There is enough marine-based ice remaining in the West Antarctic Ice Sheet today to raise global sea level by about 3.3 m (ref. 16). The example of MWP-1A serves as a reminder of the risk the ice sheet poses to the world's coasts. ■

Robert E. Kopp is in the Department of Earth and Planetary Sciences and Rutgers Energy Institute, Rutgers University, Piscataway, New Jersey 08854, USA.
e-mail: robert.kopp@rutgers.edu

HUMAN EVOLUTION

Those feet in ancient times

A fossil foot found in Ethiopia suggests that human ancestors that walked on two feet and also ably climbed trees existed until 3.4 million years ago, adding evidence for locomotor diversity during early human evolution. [SEE ARTICLE P.565](#)

DANIEL E. LIEBERMAN

The limitations of the fossil record leave ample room for debate about human origins. But most palaeoanthropologists agree that selection for bipedalism was instrumental in setting the human lineage on its separate evolutionary path from the chimpanzee lineage. And, as with any journey, it was probably sensible for our ancestors to put their best foot forward when starting out. The big question is, what kind of foot? On page 565 of this issue, Haile-Selassie and colleagues¹ present findings from a partial foot fossil which suggest that the feet of early hominins (species more closely related to humans than to chimpanzees), and hence their locomotor behaviour, were more diverse than was previously thought, and that the diversity lasted for much longer than was thought.

Human feet are remarkably different from those of apes² (Fig. 1). We have long, hefty big toes whose orientation does not diverge from that of the other toes, which are shorter and straighter than in other primates. Our feet also have a large, stable heel, for striking the ground

when we walk, and a well-developed arch that stiffens the middle of the foot and transfers the body's weight inward towards the base of the big toe, helping to push the body forward and upward at the end of stance.

Many of these distinctive features are also present in foot bones belonging to species of *Australopithecus*, a diverse genus of hominin that lived in Africa from about 4.4 million to 1.3 million years ago³. An absence of fossil feet older than those of *Australopithecus* led palaeoanthropologists to believe that human-like feet helped guide the way in human evolution, by enabling early hominins to walk effectively as bipeds, even while they retained some features that helped them to climb trees. In addition, the origin of the genus *Homo*, to which modern humans belong, was thought to have involved only minor modifications to foot anatomy, perhaps to improve our ancestors' ability to run long distances, although at the expense of climbing⁴.

Recent discoveries have made that simple narrative more complex. Most importantly, a fossil foot from *Ardipithecus ramidus*⁵, a 4.4-million-year-old species of hominin,

- Clark, P. U. *et al. Science* **325**, 710–714 (2009).
- Fairbanks, R. G. A. *Nature* **342**, 637–642 (1989).
- Deschamps, P. *et al. Nature* **483**, 559–564 (2012).
- <http://publications.iodp.org/proceedings/310/310title.htm>
- Mortlock, R. *et al. in 10th Int. Conf. Paleoceanography*, La Jolla (2010).
- Peltier, W. R. *Annu. Rev. Earth Planet. Sci.* **32**, 111–149 (2004).
- Carlson, A. E., Stoner, J. S., Donnelly, J. P. & Hillaire-Marcel, C. *Geology* **36**, 359–362 (2008).
- Clark, P. U. *et al. Paleoceanography* **11**, 563–577 (1996).
- Clark, P. U., Mitrovica, J. X., Milne, G. A. & Tamisiea, M. E. *Science* **295**, 2438–2441 (2002).
- Barker, S. *et al. Nature* **457**, 1097–1102 (2009).
- Weaver, A. J., Saenko, O. A., Clark, P. U. & Mitrovica, J. X. *Science* **299**, 1709–1713 (2003).
- Carlson, A. E. *Quat. Sci. Rev.* **28**, 1625–1630 (2009).
- Mackintosh, A. *et al. Nature Geosci.* **4**, 195–202 (2011).
- Bentley, M. J. *et al. Geology* **38**, 411–414 (2010).
- Gomez, N., Mitrovica, J. X., Huybers, P. & Clark, P. U. *Nature Geosci.* **3**, 850–853 (2010).
- Bamber, J. L., Riva, R. E. M., Vermeersen, B. L. A. & LeBrocq, A. M. *Science* **324**, 901–903 (2009).

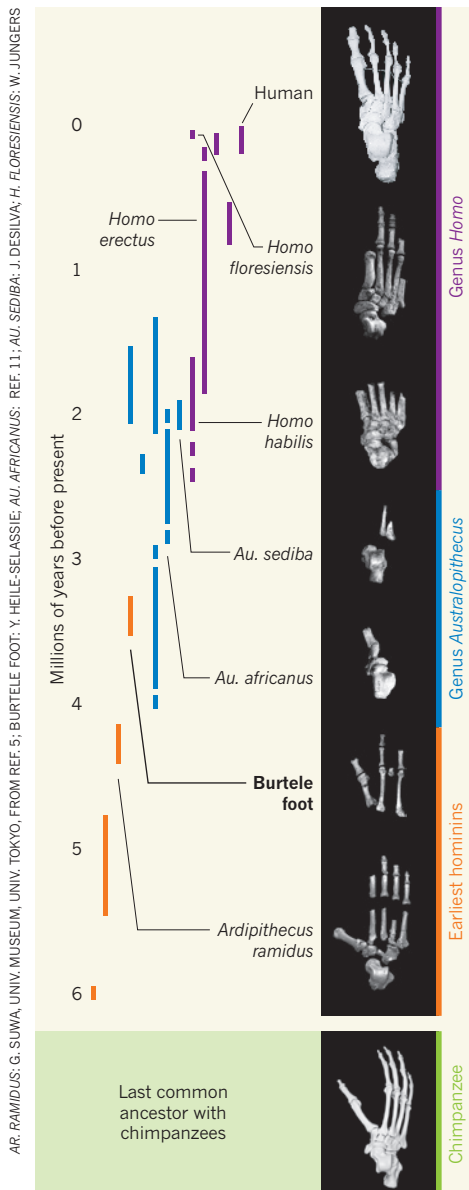


Figure 1 | Walking along the evolutionary tree.

Hominins have evolved many diverse forms of feet since diverging from their last common ancestor with chimpanzees about 6 million years ago. The early hominin species *Ardipithecus ramidus* was adapted for both walking and climbing trees⁵, but, like a chimpanzee, had a highly divergent big toe and probably used its feet more like a chimpanzee than like a modern human when it walked. Foot fossils from more recent hominin species, such as *Australopithecus sediba*, *Australopithecus africanus*, *Homo habilis* and *Homo floresiensis*, have a more complete arch than *Ar. ramidus* and a non-divergent big toe, but they were not entirely modern, retaining some adaptations for life in trees. It was probably not until *Homo erectus* that very human-like feet evolved, with a completely developed arch and a large big toe aligned with the other toes. Haile-Selassie *et al.*¹ describe bones of a fossil foot from Burtele, Ethiopia, dated to around 3.4 million years ago, which is similar to the foot of *Ar. ramidus*. This finding indicates that feet adapted to both bipedal locomotion and tree-climbing persisted for a long time in human evolution. (Foot images not to scale; some have been reflected to make them all right feet.)

shows substantial differences from the feet of *Australopithecus*. The *Ardipithecus* foot (Fig. 1) has several features suggestive of bipedalism, including evidence for a stiffened midfoot and toe joints capable of bending upward at the end of stance. But it has a very divergent and relatively short big toe, similar to that of African great apes. The foot bones also indicate that this animal placed its weight more along the lateral side of the foot when it walked, much like a chimpanzee does. The fossil's discoverers proposed⁵ that these features indicate that *Ardipithecus* was both a tree-climber and an occasional upright walker. Some researchers have argued⁶ that *Ardipithecus* was actually an ape that had independently evolved adaptations for bipedalism, whereas others, myself included, consider *Ardipithecus* to have been a hominin whose foot partly resembled an African great ape's, but with some key adaptations for bipedalism.

The foot fossil reported by Haile-Selassie and colleagues¹ is a valuable addition to the fossil record, as it extends the evidence for the existence of *Ardipithecus*-like feet by a million years. The fossil, which was discovered in fossil-rich deposits dated to 3.4 million years ago in a locality named Burtele, in the Afar region of Ethiopia, comprises eight bones, all from the front half of a single right foot (Fig. 1). In many ways, the foot is ape-like, especially resembling that of a gorilla. The big toe is short, very divergent, and apparently capable of grasping against the second toe. In addition, the toe bones are generally long and slightly curved, placing them between those of apes and hominins, although the fourth metatarsal bone is curiously long, like a monkey's.

However, the foot bears several traces of adaptation for bipedalism. Most tellingly, the ends of its metatarsal bones (other than those in the big toe) are large and spherical, and the matching phalange bones, which form joints with the metatarsals, have upwardly canted ends. These features, which are typical of later hominins (but also variably present in chimps and gorillas⁷), suggest that the Burtele foot was able to hyperextend its toes to help push off at the end of stance. Although there is no indication that the foot has a longitudinal arch, as was the case in *Australopithecus*^{3,8,9}, the tall base of its first metatarsal bone hints at the presence of a transverse arch.

Haile-Selassie and colleagues have not yet assigned the Burtele foot to a particular species, as more fossils are needed to make a secure assessment. However, the resemblance of this fossil, from 3.4 million years ago, to the 4.4-million-year-old foot of *Ar. ramidus* suggests that ardiopithecids were both climbing trees and walking in eastern Africa at the same time that *Australopithecus afarensis* was walking around in that region — sometimes leaving footprints that strongly suggest a human-like gait¹⁰. In other words, if *Ardipithecus* was a hominin (as I think it was), then

it seems that there was more diversity in hominin locomotion than was previously thought, and not all of it took place on the ground. Additional evidence for this diversity comes from foot bones of the recently discovered *Australopithecus sediba*, which lived approximately 2 million years ago in South Africa⁹. This species' fascinating foot (Fig. 1) has many adaptations for bipedalism, including an arch, but its ape-like heel and other features in its ankle suggest that it walked on an inwardly angled foot (like an ape), while retaining other adaptations for climbing trees.

Taking the next step to understanding the implications of this limb diversity for human evolution will require researchers to continue getting their feet dirty in the field and the lab. We need more fossils to determine what sorts of bodies went with these feet, and to resolve which features evolved just once and which evolved multiple times. We also need to have a better understanding of how the anatomical variations we see in hominin feet affected the different species' ability to climb, walk and run. For example, how much did a divergent big toe and keeping weight on the outside of the foot affect early hominins' ability to walk effectively? And to what extent did the more human-like foot of *Australopithecus* compromise its ability to climb trees? Whatever the answers, it is evident that hominin feet, like heads, were adaptively diverse, and that tree-climbing remained an important part of the hominin locomotive repertoire for several million years.

Human evolution is often portrayed as a triumph of bipedalism, but who among us has not occasionally regretted our species' comparative clumsiness in trees? I, for one, am pleased to know that some hominins retained feet well adapted for arboreality millions of years after we started to walk on two feet. ■

Daniel E. Lieberman is in the Department of Human Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA.
e-mail: danlieb@fas.harvard.edu

1. Haile-Selassie, Y. *et al.* *Nature* **483**, 565–569 (2012).
2. Klennerman, L. & Wood, B. *The Human Foot: A Companion to Clinical Studies* (Springer, 2006).
3. Harcourt-Smith, W. E. H. & Aiello, L. C. *J. Anat.* **204**, 403–416 (2004).
4. Bramble, D. M. & Lieberman, D. E. *Nature* **432**, 345–352 (2004).
5. Lovejoy, C. O., Latimer, B., Suwa, G., Asfaw, B. & White, T. D. *Science* **326**, 72, 72e1–72e8 (2009).
6. Wood, B. & Harrison, T. *Nature* **470**, 347–352 (2011).
7. Griffin, N. L. & Richmond, B. G. *Am. J. Phys. Anthropol.* **141**, 116–123 (2010).
8. Ward, C. V., Kimbel, W. H. & Johanson, D. C. *Science* **331**, 750–753 (2011).
9. Zipfel, B. *et al.* *Science* **333**, 1417–1420 (2011).
10. Raichlen, D. A., Gordon, A. D., Harcourt-Smith, W. E. H., Foster, A. D. & Haas, W. R. *J. PLoS ONE* **5**, e9769 (2010).
11. Clarke, R. J. & Tobias, P. V. *Science* **269**, 521–524 (1995).

Evidence against a chondritic Earth

Ian H. Campbell¹ & Hugh St C. O'Neill¹

The $^{142}\text{Nd}/^{144}\text{Nd}$ ratio of the Earth is greater than the solar ratio as inferred from chondritic meteorites, which challenges a fundamental assumption of modern geochemistry—that the composition of the silicate Earth is ‘chondritic’, meaning that it has refractory element ratios identical to those found in chondrites. The popular explanation for this and other paradoxes of mantle geochemistry, a hidden layer deep in the mantle enriched in incompatible elements, is inconsistent with the heat flux carried by mantle plumes. Either the matter from which the Earth formed was not chondritic, or the Earth has lost matter by collisional erosion in the later stages of planet formation.

The paradigm that underpins much of modern geochemistry is that the integrated chemical composition of the whole Earth should be that of the Sun, except for depletion in volatile elements, according to their volatility under the conditions of the solar nebula. Similar solar-related compositions are found in ‘chondritic’ meteorites, which are fragments of small rocky bodies that escaped the usual course of planetary differentiation into a metallic core, silicate mantle and crust. The composition of a chondritic meteorite is therefore presumed to reflect its entire parent body. Although the solar composition can be determined from spectroscopic measurements of the solar photosphere, measurement is not possible or imprecise for many elements, is model-dependent and does not give information on the isotopic make-up of the elements¹. Instead, a more complete picture of the solar composition is inferred from chemical analyses of chondrites. The compositions of the chondrites vary, with at least 27 parent bodies sampled², reflecting local differences in the solar-nebula-like density or proportions of gas to solids, or different accretion processes. The various chondrite compositions are distinguished by enrichment or depletion of refractory elements, ratio of lithophile to siderophile elements (for example, Mg/Fe), oxidation state, oxygen isotopic compositions, and their patterns of depletion of the volatile elements. No examples with volatile-element enrichment are known, except for a slight enrichment in a few of the least volatile of these elements in the highly reduced enstatite chondrites. Yet all chondrites share one distinctive compositional feature: their refractory lithophile elements (RLEs) are present in the same ratio relative to each other and to the solar composition. The RLEs are defined by two properties: they are refractory, because they condense from a gas of solar composition at temperatures higher than the main constituents of rocky planets, the magnesium silicates and iron metal; and they are lithophile, because they do not enter metal or sulphide phases, either in chondrites or into the metallic cores formed during planetary differentiation. There are 28 RLEs that are stable or have long half-lives; they include Ca and Al among the major elements, the entire suite of rare earth elements (REEs), and the radiogenic heat-producing elements U and Th.

The constant RLE ratio rule is ever challenged on several fronts: by exceptions due to terrestrial weathering or *ad hoc* effects on parent bodies such as impact brecciation, incipient melting or aqueous alteration; by the scale of heterogeneity in chondrites relative to the small sample sizes available for analysis; by improvements in the precision of chemical analysis; and by the increasing numbers of chondritic meteorites available for analysis. It is therefore difficult to quantify the precision to which the rule holds, but variations from the solar ratios reflecting whole-body chemistry that are larger than a few per cent are exceptional. The

REE pattern in the RLE-rich CV chondrite Allende is perhaps the largest well-attested deviation³. New techniques of isotopic analysis are revealing small anomalies in the isotopic make-up of heavy elements in bulk samples of chondrites, ascribed to less than perfect homogenization of different nucleosynthetic components in the solar nebula, such as Ti (ref. 4), Ni (ref. 5), Ba (refs 6 and 7) and Mo (ref. 8). This evidence challenges the conceptual basis behind the constant RLE ratio rule, but as yet at no more than the few-per-cent level already accepted. For example, although the range in Lu/Hf and Sm/Nd in unequilibrated carbonaceous, ordinary, and enstatite chondrites is as much as 7.9% and 3.5% respectively, the average Lu/Hf and Sm/Nd values for these three main classes of chondrites agree within 1% and 0.3% respectively⁹.

Although most geochemists long ago abandoned the notion that the Earth's composition mimics any particular type of chondrite¹⁰, the idea that the Earth has solar ratios of RLEs has persisted, providing the fundamental reference frame against which trace element and radiogenic isotopic ratios are compared. This reference frame has usually, if confusingly, been termed ‘chondritic’ rather than ‘solar’ in the literature, because of the history of fine-tuning the presumed solar composition to the compositions of chondrites. Emphasis has been placed on the CI chondrites, which match the solar composition within uncertainty for many elements irrespective of their chemical properties, except for the most volatile elements; but CIs are rare, and useful analyses come from just three falls, Orgueil, Ivuna and Alais¹¹, and mostly from Orgueil¹. The implicit assumption is that the bulk chemical composition of rocky bodies is established at the earliest stages in the planet-building process, of which the chondrite parent bodies are relicts. In this view, the subsequent stages by which these small chondritic bodies collide and merge to form planetary embryos and ultimately Earth-sized planets, while resulting in extensive differentiation associated with melting, does not affect the integrated whole-planet compositions: these, it is assumed, remain ‘chondritic’.

Planetary accretion

The ‘chondritic’ hypothesis for the Earth's composition is a survivor from times when the understanding of how terrestrial planets form was quite different. Current models can be traced back to Safronov, whose monograph on the subject was only translated into English in 1974 (ref. 12). Our present understanding is that planetary accretion proceeds through several stages of ever increasing average size. Initially, runaway growth is from kilometre-sized bodies (approximately the size of parent bodies of chondrite meteorites) to form planetesimal oligarchs about a thousand kilometres in diameter. This is the approximate size of the differentiated bodies that are thought to be parental to the achondrite meteorites, like Vesta (diameter 500 km). The achondrite parent bodies

¹Research School of Earth Science, Australian National University, Canberra, Australian Capital Territory 0200, Australia.

all show non-chondritic patterns of volatile loss, for example, in their Mn/Na ratios¹³. This post-nebular volatile loss can be dated using the Rb/Sr chronometer to several million years after the chondrite stage¹⁴. The last stages of planetary accretion see the assembly of these planetesimals into Moon-to-Mars-sized planetary embryos, which then merge through highly energetic collisions to form a few terrestrial planets. The giant impact that formed the Moon is the last of significance in the formation of the Earth. None of this was envisaged when the chondritic hypothesis was first advanced. Maintaining the hypothesis in its current form requires that the collisions between bodies during the several stages of accretion result in no net fractionation of RLEs, despite the bodies being already differentiated at the earliest stages, which does not seem probable.

Challenge to the chondrite paradigm

The challenge to this paradigm started with two landmark papers showing that the $^{142}\text{Nd}/^{144}\text{Nd}$ ratio of chondritic meteorites is 20 ± 5 parts per million less than rocks of terrestrial mantle origin^{15,16}. ^{142}Nd is the daughter of ^{146}Sm , with a half-life generally assumed to be 103 million years (Myr) but which could be as short as 68 Myr (ref. 17). Sm and Nd are two typical RLEs, and their isotopic relationships provide particularly powerful constraints on Earth differentiation models, because the short-lived $^{146}\text{Sm}/^{142}\text{Nd}$ system is complemented by the long-lived $^{147}\text{Sm}/^{143}\text{Nd}$ system (half-life 106 billion years, Gyr), which has long been used to constrain the sizes of Earth reservoirs. Although there is some nucleosynthetic variability in the isotopic composition of Nd among chondrites, this is unlikely to explain the difference in $^{142}\text{Nd}/^{144}\text{Nd}$ (ref. 18), which therefore requires the ratio of Sm to Nd to be about 6% above the average chondritic value^{15,16}; because of its isotopic significance, this value is known to within 0.3% (refs 9 and 19).

There are two possible interpretations¹⁵. The simplest is that the Earth is not chondritic after all, and the measured $^{142}\text{Nd}/^{144}\text{Nd}$ ratio of terrestrial samples is that of the bulk silicate Earth (BSE). Many geochemists have opted for an alternative scenario, in which the Earth's mantle underwent an early fractionation event into an early-enriched reservoir with low Sm/Nd and an early-depleted reservoir (EDR) with high Sm/Nd^{15,16,20–22}. Boyet and Carlson (refs 15 and 16) suggest that the low-Sm/Nd reservoir was an incompatible-element-enriched basaltic crust that sank to the core–mantle boundary, becoming isolated in the seismically anomalous region at the base of the mantle called D'', an irregular 200–250-km-thick layer overlying the core. The seismic properties of D'' make it unlikely to be a simple thermal boundary layer and it is interpreted to be a stable layer with a density greater than that of the overlying mantle^{23,24}. Furthermore, because the half-life of ^{146}Sm is no more than 103 Myr, the early fractionation event must have occurred well within the first 10 Myr of Solar System formation to prevent the average $\epsilon^{143}\text{Nd}$ value of the complementary EDR rising above about 10, the observed average value of the mid-ocean-ridge basalt (MORB)^{25,26}.

The $^{142}\text{Nd}/^{144}\text{Nd}$ ratios of lunar samples, however, are indistinguishable from terrestrial values^{27–29}, so the Moon and the Earth developed their $^{142}\text{Nd}/^{144}\text{Nd}$ enrichment before the Moon's formation³⁰. If the Moon formed by a giant impact, the collision would have melted and homogenized the Earth's mantle³¹, which would have destroyed any hypothetical early-enriched reservoir³⁰. Furthermore, the effect of extraction of the continental crust from the mantle on the Sm–Nd mass balance can be modelled, assuming that the BSE has a composition similar to the EDR, without a hidden reservoir³⁰.

The heat carried by mantle plumes provides an equally compelling argument against a hidden early-enriched reservoir³². Most hidden-reservoir advocates^{15,16,33,34} equate it with D''. If D'' is the hidden early-enriched reservoir, then under the chondritic assumption, a mass balance requires it to contain over 40% of the Earth's heat-producing elements (U, Th and K), which produce 9 terawatts (TW) of heat¹⁵. Regardless of whether D'' is a stable layer or forms a double-diffusive convecting layer³⁵, the heat it liberates can only be transmitted through the overlying mantle in plumes. The amount of heat transferred by mantle plumes can be estimated from the dynamic topography they

generate on the sea floor. Early estimates^{36,37} placed the heat flow carried by plumes at 3.5 TW, which has been revised³⁸ to 7 TW or 15% of the Earth's total heat flow³⁹ of 47 ± 2 TW, to take into account the mantle's subadiabatic thermal gradient. However, models of mantle heat transfer, which takes into account both the heat required to warm subducted lithosphere and the additional heat required to lift compositionally dense plumes, suggest a higher figure^{38,40–42} of 7–14 TW. From this must be subtracted the heat transferred from the core to the mantle, which⁴¹ must be at least 3–4 TW, the minimum amount of heat required to sustain the geodynamo, but is more likely to lie between 5–7 TW and 12–14 TW (ref. 38). As a consequence, if heat transfer from the core to the mantle is greater than the low estimate of 5 TW or if the heat carried by mantle plumes is less than the high estimate of 14 TW, 40% of the Earth's heat-producing elements cannot be hidden in the D'' layer.

A sudden drop in the maximum MgO content of plume-related komatiites and picrites 2.5 Gyr ago from 30–35 wt% to 18–23 wt%, which implies a temperature drop of 200 to 250 °C, has also been used to argue that D'' did not form until the end of the Archaean eon³⁵. The simplest explanation for the observed drop in MgO is that D'' formed as a stable layer about 2.5 Gyr ago, which insulated the mantle from the core. The predicted drop in plume temperatures, depending on assumptions such as whether D'' is a stable layer or formed a double-diffusive convecting layer, lies within the range 33% to 50%, which is consistent with the observed MgO drop³⁵. If D'' formed after the first 10 Myr of the Earth's evolution it cannot be responsible for the Earth's $^{142}\text{Nd}/^{144}\text{Nd}$ anomaly.

If the composition of the EDR is that of the BSE²⁰, or if it formed within 10 Myr of the Solar System^{25,26} and is unaffected by subsequent events, then the $\epsilon^{143}\text{Nd}$ value for the EDR today^{30,43} is 7.0 ± 2.0 . These values are remarkably similar to the prevalent mantle (PREMA) value of 7 ± 1 in ocean island basalts (OIBs)^{44,45}. A component with this Nd characteristic can be recognized in four flood basalts⁴³—the Baffin Island–West Greenland province, the Antarctic Karoo, the Siberian Traps and the Deccan Traps—and in two oceanic plateaus—the Kerguelen and Ontong Java plateaus³⁵. The EDR component plots within 150 Myr of the 4.5-Gyr geochron on a $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram, which suggests that it developed early in the Earth's history⁴³. It also has enriched $^3\text{He}/^4\text{He}$, showing that the source region to these plumes is less degassed than the MORB source^{30,43}.

Torsvik *et al.* suggest that most plume-related basalts that erupted over the last 320 Myr were above one of two large low-shear-wave-velocity provinces, which form part of D'' and make up about 2% of the mantle's mass⁴⁶. However, this primitive component in plumes may be much more extensive^{35,45}. The isotopic trend in a number of ocean island suites, produced by melting plume tails, converges towards a common point called FOZO⁴⁴. A remarkably similar component was identified in the two oceanic plateaus that are free from contamination by continental crust, the Ontong Java and the Caribbean plateaus^{35,43}. Basalts from these plateaus are characterized by flat REE patterns and in this respect they are dissimilar to OIBs and MORBs but are isotopically similar to the EDR recognized in flood basalts⁴³. Archaean basalts associated with komatiites, whose Nb/U ratios indicate that they are free of crustal contamination^{47–50}, commonly have flat REE patterns^{48–50}, and basalts from Kambalda in Western Australia have an $\epsilon^{143}\text{Nd}$ value of 3 (ref. 47), which lies on the EDR growth curve for Nd at 2.7 Gyr ago. Jackson and Carlson⁴³ interpret the EDR component to originate from the boundary layer source of plumes, whereas Campbell and Griffith³⁵ suggest that it is lower mantle that was entrained into plumes during their ascent. The first interpretation allows the early-enriched reservoir to be a minor component in plumes, whereas the second requires it to be a dominant component in the lower mantle. Because this difference is critical to the debate on whether the Earth has chondritic RLE ratios, we now summarize the basis for the second interpretation^{35,51}.

Plumes must originate from a thermal boundary layer; in the Earth, only the core–mantle boundary has the properties required to sustain the temperature drop implied by plume activity on geological timescales⁵². A

new plume has a large head, which is followed by a smaller tail (Fig. 1). As the head rises through the mantle it heats the adjacent mantle, lowering its density so that it is swept into the plume head by its recirculating motion. The plume head is therefore a mixture of material from the hot boundary-layer source of the plume (dark material in Fig. 1) and cooler entrained material (light material in Fig. 1). This entrained material, which makes up a large fraction of the head, comes from the lower half of the lower mantle^{51,52}. When a plume head reaches the top of the mantle it melts to produce a continental flood basalt or an oceanic plateau^{52,53}. The first material to reach the top of the mantle and undergo decompressional melting is the hot mantle that originated from the boundary layer above the core (Fig. 1), which melts to produce picrites or komatiites. Later, when the head flattens against the overlying lithosphere, the cooler entrained material from the lower mantle, which is calculated to be three times as thick as the hot material from the boundary layer⁵⁴, enters the melting zone. Production of high-temperature picrites or komatiites should therefore be followed by lower-temperature basalts formed by melting a mixture of boundary layer and entrained lower-mantle with the latter becoming increasingly important with time as the plume head continues to rise and flatten. It is this primitive entrained lower-mantle component that has been identified in the Ontong Java and Caribbean plateaus^{35,43,51} and in several flood basalt provinces⁴³.

Lower-mantle material is also entrained along the side of rising plume tails by viscous drag (Fig. 1). As a consequence, when plume tails melt to produce a chain of ocean islands, which progressively increase in age away from the current position of the plume, the basalts that make up the islands may contain a component from the lower mantle. The existence of FOZO as a component in plume tail basalts is therefore consistent with the interpretation⁴⁵ that it was entrained from the lower mantle during ascent of a plume tail (Fig. 1). Its presence in all or most plume tails provides further support that FOZO is a major component in the lower mantle.

FOZO⁴⁵, EDR¹⁵, SCHEM³⁰ (super-chondritic Earth model), and NCPM⁴³ (non-chondritic primitive mantle) are therefore all the same: the primordial component of the BSE. The evidence from oceanic plateaus, flood basalts and OIBs is that this is a major component in the lower

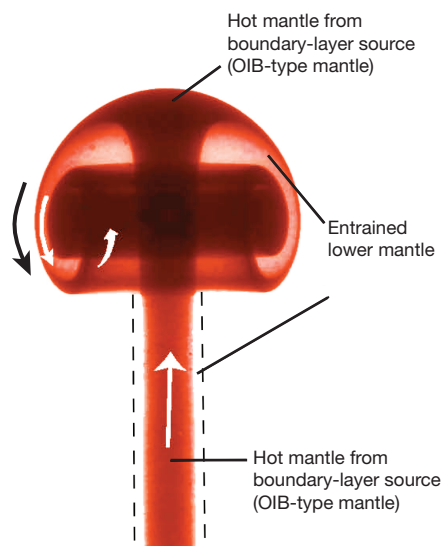


Figure 1 | Laboratory model of a mantle plume. The dark-coloured fluid is from the hot boundary-layer source of the plume, whereas the light material is cooler overlying fluid that was entrained into the rising plume. In the case of the mantle the dark-coloured material is from the thermal boundary layer above the core, whereas the light material is entrained lower mantle (after ref. 54). The entrained material makes up a large fraction of the plume head⁵⁴ and comes mainly from near the bottom of the lower mantle but above the boundary-layer source of the plume. The critical parameter in the present context is the thickness of the upper entrained layer (light material near the top of the plume head), which is about three times the thickness of layer at the top of the head (dark material) that originates from the boundary-layer source of the plume⁵⁴.

mantle. The high $^3\text{He}/^4\text{He}$ ratios recognized in the EDR component are consistent with this interpretation³⁰. The displacement of most oceanic plateaus and flood basalts, with $\epsilon^{143}\text{Nd}$ between 5 and 9, slightly to the right of the geochron on a plot of $^{207}\text{Pb}/^{204}\text{Pb}$ against $^{206}\text{Pb}/^{204}\text{Pb}$, shows that the lower mantle, as sampled by Phanerozoic plumes, includes some recycled oceanic crust⁴³. Basalts from Malaita Island, part of the Ontong Java Plateau, have an average Nb/U of 42, which is between that of the average for both the OIB- and MORB-type mantles (47) and the BSE value of 34, indicating that some continental crust has also been extracted from the lower mantle⁵¹.

The picture of mantle convection that is emerging is one in which the upper mantle differentiates into harzburgite and basalt at mid-ocean ridges, which are subducted to the thermal boundary layer above the core and returned to the upper mantle in plumes, mixed in varying proportions, with perhaps a small sedimentary component. It appears that the bulk of the lower mantle is largely bypassed by this process, so that it is less degassed, has had less subducted basalt mixed into it, and less continental crust extracted from it, compared to OIB- and MORB-source mantles.

Volatile elements

Although the comparison of Earth's composition with chondrites has focused on the RLEs, the differences in the abundances of the moderately volatile elements between the Earth and chondrites provide another line of evidence that chemical fractionations, which occurred in the latter stages of accretion, moved the final compositions of terrestrial planets away from the compositions encompassed by chondritic meteorites^{13,14}. The moderately volatile elements are those calculated to condense from the solar nebula after magnesian silicates and Fe metal, but at temperatures above the ice-forming elements (H, C, N and the noble gases). Many moderately volatile elements are also siderophile and were depleted in the BSE by partitioning into the core, making their bulk Earth abundances inaccessible, but among the lithophile elements are the alkali metals, the halogens and boron. For the other potentially siderophile elements like Zn and In, their BSE abundances provide useful constraints on minimum Earth abundances. The pattern of depletion of the moderately siderophile elements is unlike that found in any chondrite group (Fig. 2). In the chondrites, depletion correlates with calculated condensation temperatures, but in the Earth, elements that are very highly incompatible during partial melting, namely the heavy halogens and Cs, are much more depleted than the nearly compatible Zn and In. This is highlighted by Zn/Cl ratios, which are an order of magnitude greater in the BSE than in any chondrite group (Fig. 2). The dependence of volatility on incompatibility suggests volatilization from early-formed crusts during the latter stages of accretion¹³. In the case of the halogens, their volatilities are enhanced by relatively oxidizing conditions that prevailed after dispersion of the H-rich solar nebula. Post-nebular oxidizing conditions are reflected in the ubiquitous fractionation of Na from Mn in small-differentiated planetary bodies¹³, although this particular fractionation is not seen in the nearly chondritic Na/Mn ratios of the BSE, perhaps because some Mn has also been partitioned into the core. Attempts to define simple volatile-element depletion trends in the BSE have invariably omitted key elements from consideration, like the heavy halogens.

For elements that are both volatile and siderophile, the complexities of the volatile-element depletion in the Earth, as indicated by the full pattern (Fig. 2), prevents us from assigning how much of an element's loss is due to volatility and how much is due to partitioning into the core. This is particularly vexing for Pb. Was the extent of Pb depletion by volatility in the material that formed the Earth like that of Zn and In, or was it like Cl and the other heavy halogens, or was it somewhere in between? On the $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram, it is well established that both the mantle and crust of the Earth plot well to the right of the 4.57-Gyr geochron, which, neglecting the 'hidden reservoir' explanations, is usually taken to imply loss of Pb relative to parental U sometime later, at about 4.45 Gyr ago⁵⁵. Because feasible hypotheses can be constructed to argue for Pb loss by both mechanisms at several stages

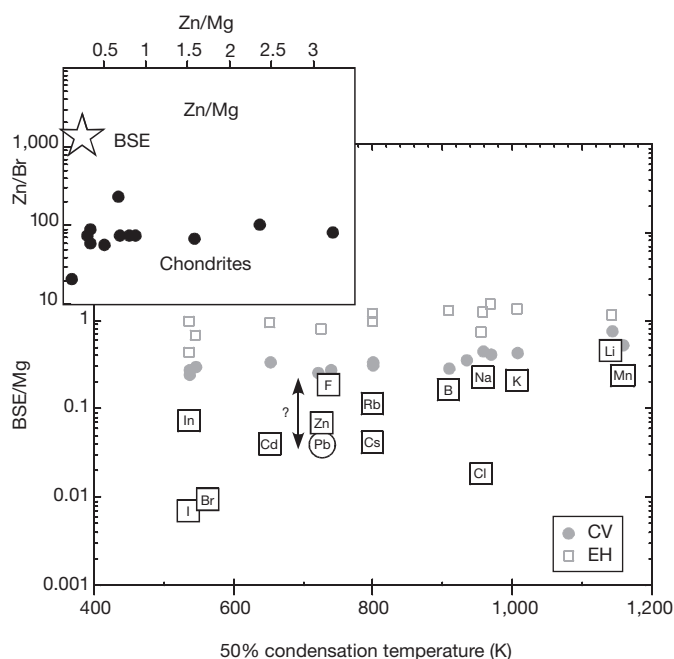


Figure 2 | The pattern of volatile element depletion in the BSE for lithophile elements compared to CV carbonaceous chondrites and EH enstatite chondrites. CV carbonaceous chondrites are the most volatile-depleted of the chondrites and EH enstatite chondrites are a class of chondrites sometimes considered to have affinities to the Earth because of their stable isotope ratios. Elements are normalized to CI abundances and Mg. The calculated 50% condensation temperatures of elements for the solar nebula are from ref. 10. Some elements in the BSE may have been additionally depleted by core formation (for example, Zn, In and Pb), in which case their depletions due to volatility alone will be overestimated. The chondrites form smooth depletion trends with calculated condensation temperatures, but the BSE is not only more depleted in moderately volatile elements than any known chondrite, but the pattern of depletion is qualitatively different, probably owing to post-nebula volatile loss under more oxidizing conditions¹³. This is shown, for example, by the BSE Zn/Br ratios (inset), which are about an order of magnitude greater than that found in any class of unmetamorphosed chondrites (black circles). The lack of any clear volatility trend in the BSE means that it is not at present possible to constrain how much Pb, if any, was partitioned into the Earth's core, making the interpretation of Pb isotopic systematics in terms of Earth accretion and core formation uncertain. Meteorite abundances are from ref. 66, and BSE abundances from ref. 13, which was derived under the 'chondritic assumption'. A non-chondritic BSE would have lower abundances of B, K, Rb, Cs, Cl, Br and I but not Mn, Na, F, Zn or In (ref. 12), increasing the discrepancy with the chondritic trends.

of the planet-building process, the significance of this age information remains ambiguous.

Likewise, the interpretation of other early-Earth geochronometers is also affected if the Earth has a non-chondritic composition. For example, short-lived ^{182}Hf (half-life 9 Myr) decays to ^{182}W , providing a chronometer with which to constrain the duration of core formation, because W is a moderately siderophile element that partitioned incompletely into the core. Evaluating the time significance of this chronometer for the Earth depends on knowing the W/Hf ratio in the BSE⁵⁶. The W content of the BSE has been estimated by noting that W/Th (or W/Ba) ratios remain constant in igneous processes, hence $\text{W/Hf} = (\text{W/Th}) \times (\text{Th/Hf})$. Because Hf and Th (or Ba) are both RLEs, it is then assumed that their ratio is the same as in chondrites. But the non-chondritic Earth model of ref. 13 predicts that Th/Hf would be only around 70% of the chondritic ratio; for a simple two-stage model of core formation, this would increase the calculated time from about 30 to about 35 Myr.

Alternative hypotheses

There are two classes of explanation for the Earth not being chondritic. Most simply, the compositions of chondrites may not reflect that of the

Solar System precisely enough to deduce detailed element ratios for RLE. It needs to be remembered, however, that the average Sm/Nd ratio of the three main classes of chondrites agrees to within 0.3%.

Alternatively, the Earth could have been assembled from initially chondritic material that was then modified during the subsequent stages of the planet-building process by collisional erosion^{13,57,58}. Current estimates of the Earth's Fe/Mg ratio are consistent with about 10% of its silicate part having been lost by this mechanism relative to its Fe-rich metallic core. The meteorite record attests to the differentiation of small rocky bodies into metal and silicate being inevitably associated with partial melting and hence also the formation of an incompatible-element-enriched crust. If material from these crusts were preferentially lost during the collisions, it would deplete the Earth systematically in incompatible elements according to their incompatibility (Fig. 3). One weakness of the hypothesis is that it implies the loss of incompatible-element-enriched material to space that no class of meteorite has sampled. However, no meteorites sample the moderately volatile elements missing from all chondrites (apart from the CIs) and from the achondrites and terrestrial planets. Did the gravitation field of the Sun or Jupiter capture this missing material?

The pattern of depletion caused by preferential collisional erosion is geochemical rather than cosmochemical, and its effect on the composition of the Earth's mantle is essentially the same as that which subsequently occurred throughout the Earth's history by crust formation. Detecting the effects of collisional erosion therefore depends on observations that can sum all the reservoirs in the BSE to see whether they add up to chondritic ratios of RLEs. The Sm/Nd ratio provides the most compelling evidence, but once the idea of a non-chondritic Earth is allowed, the resolution of other so-called geochemical paradoxes becomes achievable.

Geochemists are fascinated by the many paradoxes of the Earth's mantle, which are summarized in Table 1. All of these paradoxes are predicated on geochemistry's most fundamental paradigm; that the Earth was produced by the accretion of meteorites with the same ratios of RLEs as in chondrites. Most of these paradoxes disappear if this assumption is relaxed, but one existing paradox becomes worse: the low value of the Earth's Urey ratio. The Urey ratio is the Earth's radiogenic heat production divided by its surface heat flux, which—under the assumption of BSE U, Th and K concentrations given by the chondritic hypothesis—is about 0.5. The difference must be accounted for by secular cooling. Collisional erosion lowers the heat-producing element content of the BSE by up to

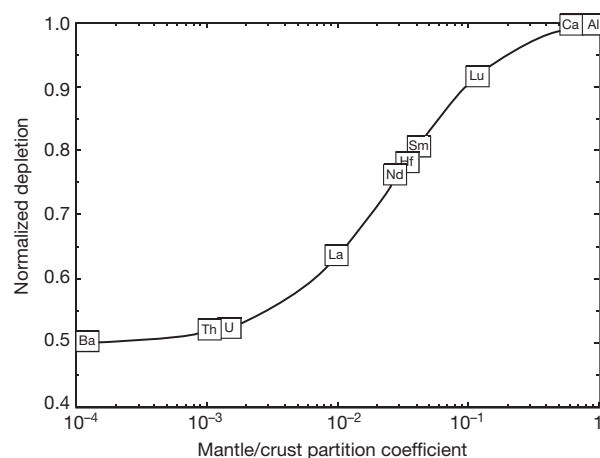


Figure 3 | Depletion of some RLEs in the BSE by preferential collisional erosion of early-formed basaltic crust during accretion of planetesimals and planetary embryos. The figure is based on the model of ref. 13, assuming three constraints: (1) loss of silicate (crusts plus mantles) relative to metallic cores is 10%; (2) the most incompatible RLEs (here represented by Ba) are depleted to 50% of their chondritic abundance; and (3) Sm/Nd is 6% above the chondritic ratio. The partition coefficients during crust formation are from ref. 67 for the production of oceanic crust.

Table 1 | The geochemical paradoxes of the mantle

Paradox	Chondritic solution	Non-chondritic solution
The $^{142}\text{Nd}/^{144}\text{Nd}$ ratio of chondritic meteorites is 20 ± 5 parts per million less than that of rocks of terrestrial mantle origin.	A low-Sm/Nd hidden reservoir became isolated from the convecting mantle within 10 Myr of the Earth's formation ^{15,61} .	The Sm/Nd ratio of the primitive Earth was about 6% above the chondritic value ¹³ .
Earth's oldest rocks show evidence of being derived from a mantle with positive ϵ_{Nd} and ϵ_{Hf} before the formation of the first preserved continental crust.	Extensive continental crust formed before the first preserved continental crust and was recycled through the mantle ⁶² or there is a hidden basaltic low-Sm/Nd reservoir ¹⁵ .	The Sm/Nd ratio of the primitive Earth was about 6% above the chondritic value ¹³ .
The Ar concentration in the mantle is about half the value predicted from the chondritic model ⁶³ .	Only half of the mantle is degassed ⁶³ .	The collisional erosion hypothesis ¹³ predicts a K content of the mantle appreciably below that expected from the chondritic model. Alternatively, the Earth is not chondritic ¹⁶ .
Nb/Ta and Nb/La values of both continental crust and depleted mantle lie below (Nb/Ta) and above (Nb/La) the primitive mantle values of 17.5 for Nb/Ta and 0.9 for Nb/La.	Hidden reservoir enriched in Nb, Ta and Nb with super-chondritic Nb/Ta and sub-chondritic Nb/La ⁶⁴ .	The Nb/Ta and La/Nb values of the primitive mantle lie between those of the depleted mantle (15.5 and 1.2) and the continental crust (12.5 and 2.2).
^4He production in oceans is less than that predicted from observed heat flow and about half that predicted from chondritic Earth model.	^4He stored in lower mantle that is separated by a boundary layer that transmits heat but not ^4He (ref. 65).	Collisional erosion model predicts the Th–U content of the BSE to be about half the chondritic value ¹³ .

half¹³, which halves the already low Urey ratio, implying unlikely cooling rates extrapolated over geological time. Perhaps the Earth is currently in a phase of abnormally fast ocean crust formation and subduction⁵⁹.

It is apparent that the only reliable way of determining the composition of the Earth is by sampling the Earth itself. As argued in this study, the heads of mantle plumes entrain primitive lower mantle. By studying basalts produced by melting this material, especially Archaean basalts associated with komatiites, provided they are not affected by crustal contamination, we are sampling basalts derived from the Earth's earliest and most primitive mantle. It may also be possible to obtain the integrated BSE composition of two of the RLEs most susceptible to collisional erosion—Th and U—from the geoneutrino flux⁶⁰. It would then be possible to see whether the ratios of these elements with other RLEs, such as Ca and Al, were indeed within the range found in chondritic meteorites or that predicted by collisional erosion.

- Lodders, K., Palme, H. & Gail, H.-P. in *Landolt-Börnstein New Series* (ed. Trümper, J. E.) Vol. VI/4B, Ch. 4.4 560–630 (Springer, 2009).
 - Meibom, A. & Clark, B. E. Evidence for the insignificance of ordinary chondritic material in the asteroid belt. *Meteorit. Planet. Sci.* **34**, 7–24 (1999).
 - Pourmand, A., Dauphas, N. & Ireland, T. J. A novel extraction chromatography and MC-ICP-MS technique for rapid analysis of REE, Sc and Y: revising CI-chondrite and Post-Archaean Australian Shale (PAAS) abundances. *Chem. Geol.* <http://dx.doi.org/10.1016/j.chemgeo.2011.08.011> (in the press).
 - Trinquier, A. *et al.* Origin of nucleosynthetic isotope heterogeneity in the solar protoplanetary disk. *Science* **324**, 374–376 (2009).
 - Regelous, M., Elliott, T. & Coath, C. D. Nickel isotope heterogeneity in the early Solar System. *Earth Planet. Sci. Lett.* **272**, 330–338 (2008).
 - Andreasen, R. & Sharma, M. Mixing and homogenization in the early solar system: clues for Sr, Ba, Nd, and Sr isotopes in meteorites. *Astrophys. J.* **665**, 874–883 (2007).
 - Carlson, R. W., Boyet, M. & Horan, M. Chondritic barium, neodymium, and samarium isotopic heterogeneity and early Earth differentiation. *Science* **316**, 1175–1178 (2007).
 - Burkhardt, C. *et al.* Molybdenum isotope anomalies in meteorites: constraints on solar nebula and origin of the Earth. *Earth Planet. Sci. Lett.* **312**, 390–400 (2011).
 - Bouvier, A., Vervoort, J. D. & Patchett, P. J. The Lu–Hf and Sm–Nd isotopic composition of CHUR: constraints from unequilibrated chondrites and implications for the bulk composition of terrestrial planets. *Earth Planet. Sci. Lett.* **273**, 48–57 (2008).
 - Drake, M. J. & Richter, K. Determining the composition of the Earth. *Nature* **416**, 39–44 (2002).
 - Lodders, K. Solar system abundances and condensation temperatures of the elements. *Astrophys. J.* **591**, 1220–1247 (2003).
 - Safronov, V. S. *Evolution of the Protoplanetary Cloud and Formation of the Earth and the Planets* (Israel Program for Scientific Translations, 1972).
 - O'Neill, H., St. C., & Palme, H. Collisional erosion and the non-chondritic composition of the terrestrial planets. *Phil. Trans. R. Soc. A* **366**, 4205–4238 (2008).
- This paper discusses how the compositions of terrestrial planets may change during the later stages of their accretion.**
- Halliday, A. N. & Porcelli, D. In search of lost planets—the paleocosmochemistry of the inner solar system. *Earth Planet. Sci. Lett.* **192**, 545–559 (2001).
 - Boyet, M. & Carlson, R. W. ^{142}Nd evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science* **309**, 576–581 (2005).

Here the $^{142}\text{Nd}/^{144}\text{Nd}$ for chondritic meteorites is shown to be 20 parts per million less than that of most terrestrial rocks.

- Boyet, M. & Carlson, R. W. A new geochemical model for the Earth's mantle inferred from ^{146}Sm – ^{142}Nd systematics. *Earth Planet. Sci. Lett.* **250**, 254–268 (2006).
- This paper discusses the implications of excess $^{142}\text{Nd}/^{144}\text{Nd}$ in terrestrial samples in terms of a non-chondritic mantle and a hidden reservoir.**
- Kinoshita, N. *et al.* Geocosmochronometer ^{146}Sm : a revised half-life value. *Mineral. Mag. (Goldschmidt Conf. Abstr.)* **75**, 1191 (2011).
 - Qin, L.-P., Carlson, R. W., Alexander, C. M. & O'Dr. Correlated nucleosynthetic isotopic variability in Cr, Sr, Ba, Sm, Nd and Hf in Murchison and QUE 97008. *Geochim. Cosmochim. Acta* **75**, 7806–7828 (2011).
 - Patchett, P. J., Vervoort, J. D., Söderlund, U. & Salters, V. J. M. Lu–Hf and Sm–Nd isotopic systematics in chondrites and their constraints on the Lu–Hf properties of the Earth. *Earth Planet. Sci. Lett.* **222**, 29–41 (2004).
 - Andreasen, R., Sharma, M., Subbarao, K. V. & Viladkar, S. G. Where on Earth is the enriched Hadean reservoir? *Earth Planet. Sci. Lett.* **266**, 14–28 10.1016/j.epsl.2007.10.009 (2008).
 - Caro, G., Bourdon, B., Birk, J.-L. & Moorbath, S. High-precision $^{142}\text{Nd}/^{144}\text{Nd}$ measurements in terrestrial rocks: constraints on the early differentiation of the Earth's mantle. *Geochim. Cosmochim. Acta* **70**, 164–191 (2006).
 - Labrosse, S., Herlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450**, 866–869 (2007).
 - Davies, G. F. Mantle plume, mantle convection and hotspot chemistry. *Earth Planet. Sci. Lett.* **99**, 94–109 (1990).
 - Loper, D. & Lay, T. The core-mantle boundary region. *J. Geophys. Res.* **100**, 6397–6420 (1995).
 - Korenaga, J. A method to estimate the composition of the bulk silicate Earth in the presence of a hidden geochemical reservoir. *Geochim. Cosmochim. Acta* **73**, 6952–6964 (2009).
 - Bourdon, B., Touboul, M., Caro, G. & Kleine, T. Early differentiation of the Earth and the Moon. *Phil. Trans. R. Soc. A* **366**, 4105–4128 (2008).
 - Nyquist, L. E. *et al.* ^{146}Sm – ^{142}Nd formation interval for the lunar mantle. *Geochim. Cosmochim. Acta* **59** (13), 2817–2837 (1995).
 - Boyet, M. & Carlson, R. W. A highly depleted moon or a nonmagma ocean origin for the lunar crust? *Earth Planet. Sci. Lett.* **262**, 505–516 (2007).
 - Brandon, A. D. *et al.* Re-evaluating $^{142}\text{Nd}/^{144}\text{Nd}$ in lunar mare basalts with implications for the early evolution and bulk Sm/Nd of the Moon. *Geochim. Cosmochim. Acta* **73**, 6421–6445 (2009).
 - Caro, G. & Bourdon, B. Non-chondritic Sm/Nd ratio in the terrestrial planets: consequences for the geochemical evolution of the mantle-crust system. *Geochim. Cosmochim. Acta* **74**, 3333–3349 (2010).
 - Tonks, W. B. & Melosh, H. J. Magma ocean formation due to giant impacts. *J. Geophys. Res.* **98**, 5319–5333 (1993).
 - Campbell, I. H. The identification of old mantle plumes. *Geol. Soc. Am. Spec. Publ.* **352**, 5–21 (2001).
 - Tolstikhin, I. N., Kramers, J. D. & Hofmann, A. W. A chemical Earth model with whole mantle convection: the importance of a core-mantle boundary layer (D') and its early formation. *Chem. Geol.* **226**, 79–99 (2006).
 - Labrosse, S., Herlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450**, 866–869 (2007).
 - Campbell, I. H. & Griffiths, R. W. The changing nature of mantle hotspots through time: implications for the geochemical evolution of the mantle. *J. Geol.* **100**, 497–523 (1992).
- This paper uses the melting products of mantle plumes in conjunction with knowledge of their dynamics to determine the chemical structure of the mantle and shows how it has evolved with time.**
- Davies, G. F. Ocean bathymetry and mantle convection. 1. Large-scale flow and hotspots. *J. Geophys. Res.* **93**, 10467–10480 (1988).
 - Sleep, N. H. Hotspots and mantle plume: some phenomenology. *J. Geophys. Res.* **95**, 6715–6736 (1990).

38. Davies, G. R. Mantle regulation of core cooling: a geodynamo without core radioactivity. *Phys. Earth Planet. Inter.* **160**, 215–229 (2007).
39. Davies, J. H. & Davies, D. R. Earth's surface heat flux. *Solid Earth*. **1**, 5–24 (2010).
40. Zhong, S. Constraints on thermochemical convection of the mantle from plume heat flux, plume excess temperature, and upper mantle temperature. *J. Geophys. Res.* **111**, B04409 (2006).
41. Lay, T., Hernlund, J. & Buffett, B. A. Core-mantle boundary heat flow. *Nat. Geosci.* **1**, 25–32 (2008).
42. Anderson, O. L. The power balance at the core-mantle boundary. *Phys. Earth Planet. Inter.* **131**, 1–17 (2002).
43. Jackson, M. G. & Carlson, R. W. An ancient recipe for flood-basalt genesis. *Nature*, (2011).
This paper demonstrates the existence of an old and primitive component in mantle plumes.
44. Zindler, A. & Hart, S. R. Chemical geodynamics. *Annu. Rev. Earth Planet. Sci.* **14**, 493–571 (1986).
45. Hart, S. R., Hauri, H. E., Oschmann, L. A. & Whitehead, J. A. Mantle plumes and entrainment: isotopic evidence. *Science* **256**, 517–520 (1992).
46. Torsvik, T. H., Burke, K., Steinberger, B., Webb, S. J. & Ashwal, L. D. Diamonds sampled by plumes from the core–mantle boundary. *Nature* **466**, 352–355 (2010).
47. Leshar, C. M. & Arndt, N. T. REE and Nd isotope geochemistry, petrogenesis and volcanic evolution of contaminated komatiites at Kambalda, Western Australia. *Lithos* **34**, 127–157 (1995).
48. Hollings, P., Wyman, D. & Kerrich, R. Komatiite-basalt-rhyolite volcanic associations in Northern Superior Province greenstone belts: significance of plume-arc interaction in the generation of the proto continental Superior Province. *Lithos* **46**, 137–161 (1999).
49. Hollings, P. & Kerrich, R. Trace element systematics of ultramafic and mafic volcanic rocks from 3 Ga North Caribou greenstone belt, northwest Superior Province. *Precamb. Res.* **93**, 257–279 (1999).
50. Kerrich, R., Polat, A., Wyman, D. & Hollings, P. Trace element systematics of Mg- to Fe-tholeiitic basalt suites of the Superior Province: implications for Archaean mantle reservoirs and greenstone belt genesis. *Lithos* **46**, 163–187 (1999).
51. Campbell, I. H. in *The Earth's Mantle: Composition, Structure and Evolution* 259–310 (Cambridge University Press, 1998).
52. Campbell, I. H. & Griffiths, R. W. Implications of mantle plume structure for the evolution of flood basalts. *Earth Planet. Sci. Lett.* **99**, 79–93 (1990).
53. Richards, M. A., Duncan, R. A. & Courtillot, V. E. Flood basalts and hot-spot tracks—plume heads and tails. *Science* **246**, 103–107 (1989).
54. Griffiths, R. W. & Campbell, I. H. Stirring and structure in mantle starting plumes. *Earth Planet. Sci. Lett.* **99**, 66–78 (1990).
55. Galer, S. J. G. & Goldstein, S. J. in *Earth Processes: Reading the Isotopic Code* (eds Basu, A. & Hart, S.) 75–98 (Geophysical Monograph 95, American Geophysical Union, 1996).
56. Kleine, T. *et al.* Hf-W chronology of the accretion and early evolution of asteroids and terrestrial planets. *Geochim. Cosmochim. Acta* **73**, 5150–5188 (2009).
57. Palme, H., O'Neill, H., St. C. & Benz, W. Evidence for collisional erosion of the Earth. *Geochim. Cosmochim. Acta* **67**, A372 (2003).
58. Caro, G., Bourdon, B., Halliday, A. N. & Quitté, G. Super-chondritic Sm/Nd ratios in Mars, the Earth and the Moon. *Nature* **452**, 336–339 (2008).
59. White, W. M. & Morgan, J. P. Implication of a non-chondritic Earth for terrestrial heat production and geodynamics. *Mineral. Mag. (Goldschmidt Conf. Abstr.)* **75**, 1191 (2011).
60. The KamLAND Collaboration. Partial radiogenic heat model for Earth revealed by geoneutrino measurements. *Nature Geosci.* **4**, 647–651 (2011).
61. Caro, G. Early silicate Earth differentiation. *Annu. Rev. Earth Planet. Sci.* **39**, 31–58 (2011).
This is a comprehensive review of Nd isotopes.
62. Armstrong, R. L. The persistent myth of crustal growth. *Aust. J. Earth Sci.* **38**, 613–630 (1991).
63. Allègre, C. J., Staudacher, T., Sarda, P. & Kurz, M. Constraints on the evolution of Earth's mantle from rare-gas systematics. *Nature* **303**, 762–766 (1983).
64. Barth, M. G., McDonough, W. F. & Rudnick, R. L. Tracking the budget of Nb and Ta in the continental crust. *Chem. Geol.* **165**, 197–213 (2000).
65. O'Nions, R. K. & Oxburgh, E. R. Heat and helium in the Earth. *Nature* **306**, 429–431 (1983).
66. Wasson, J. T. & Kallemyn, G. W. Composition of chondrites. *Phil. Trans. R. Soc.* **A325**, 535–544 (1988).
67. Workman, R. K. & Hart, S. R. Major and trace element composition of the depleted MORB mantle (DMM). *Earth Planet. Sci. Lett.* **231**, 53–72 (2005).

Acknowledgements We thank Y. Amelin and R. Taylor for their comments on an early draft of the paper, C. Allen for preparing Fig. 1 and R. Carlson and B. Bourdon for their detailed and constructive reviews.

Author Contributions Both authors contributed equally to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to I.H.C. (Ian.Campbell@anu.edu.au).

Ice-sheet collapse and sea-level rise at the Bølling warming 14,600 years ago

Pierre Deschamps¹, Nicolas Durand¹, Edouard Bard¹, Bruno Hamelin¹, Gilbert Camoin¹, Alexander L. Thomas², Gideon M. Henderson², Jun'ichi Okuno^{3,4} & Yusuke Yokoyama^{3,5}

Past sea-level records provide invaluable information about the response of ice sheets to climate forcing. Some such records suggest that the last deglaciation was punctuated by a dramatic period of sea-level rise, of about 20 metres, in less than 500 years. Controversy about the amplitude and timing of this meltwater pulse (MWP-1A) has, however, led to uncertainty about the source of the melt water and its temporal and causal relationships with the abrupt climate changes of the deglaciation. Here we show that MWP-1A started no earlier than 14,650 years ago and ended before 14,310 years ago, making it coeval with the Bølling warming. Our results, based on corals drilled offshore from Tahiti during Integrated Ocean Drilling Project Expedition 310, reveal that the increase in sea level at Tahiti was between 12 and 22 metres, with a most probable value between 14 and 18 metres, establishing a significant meltwater contribution from the Southern Hemisphere. This implies that the rate of eustatic sea-level rise exceeded 40 millimetres per year during MWP-1A.

Although dynamic responses of the Greenland and Antarctic ice sheets to climate forcing may already be contributing to present-day sea-level rise¹, projections of sea-level change for the twenty-first century do not fully include potential changes in ice dynamics². As acknowledged by the IPCC³, the vulnerability of Greenland and Antarctica to ongoing warming and related discharge feedbacks remains a major source of uncertainty in projected sea-level rise⁴. Reconstructions of past sea-level changes have provided evidence for large-amplitude and rapid discharges of fresh water from continental ice sheets. Several sea-level records suggest that the glacioeustatic rise following the Last Glacial Maximum (LGM) was characterized by brief periods of extremely rapid sea-level rise^{5–10}. These short-term events, referred to as meltwater pulses, probably disturbed oceanic thermohaline circulation and global climate during the last deglaciation^{11,12}. The exact chronology, origin and consequences of these ice-sheet melting episodes remain unclear. But understanding these episodes is of the utmost importance when considering current uncertainty surrounding potential collapse of large ice sheets in response to recent climate change¹³.

The most extreme deglacial event, MWP-1A, was initially identified in the coral-based sea-level record from Barbados⁵, where a sea-level rise of ~20 m was inferred between 14,100 and 13,600 years before present (14.1–13.6 kyr BP; from here on, all ages are given as kyr before present (BP), where 'present' refers to AD 1950)⁶. However, this event remains mysterious. Several records bear witness to its occurrence^{8,14,15}, although no broad agreement has emerged regarding its timing. Because of this lack of consensus, the temporal relationship between MWP-1A and abrupt (millennial-timescale) climatic events that punctuated the last deglaciation are the subject of considerable debate^{12,16}. Additionally, the location(s) of melting ice responsible for this prominent feature of the last deglaciation remains elusive¹⁷.

Two conflicting scenarios have been proposed to link the timing and source(s) of MWP-1A to the climatic history of the last deglaciation. On the basis of the Barbados record's chronology^{5,6}, it was initially argued that this episode of rapid sea-level rise was caused by a partial melting of Northern Hemisphere ice sheets (NHIS)^{5,18,19}. This

'Northern' scenario was consistent with results from a coupled ocean–atmosphere general circulation model (GCM), in which massive freshwater input to the North Atlantic would result in a weakening of the Atlantic meridional overturning circulation (AMOC) and, through the reduction of deepwater formation in the Nordic Seas, the rapid cooling of the Northern Hemisphere¹¹. In this scenario, MWP-1A may have initiated the Older Dryas cold event that abruptly ended the Bølling warming about 14.1 kyr ago^{14,16}.

In contrast, an alternative scenario points towards an Antarctic ice sheet (AIS) as the source of MWP-1A^{17,20} and suggests a causative coupling between MWP-1A and the Bølling warm period²¹. This 'Southern' scenario suggests that MWP-1A coincided with an intensification of the thermohaline circulation at the onset of the Bølling warm period²², rather than with a slowdown during the following cold event as predicted by the 'Northern' scenario. The 'Southern' scenario was supported by output from a GCM model of intermediate complexity showing that an MWP-1A originating from the West Antarctica Ice Sheet (WAIS) may have triggered sudden reactivation of the AMOC to lead to the Bølling warming¹². Although still contentious, this scenario solves the apparent conundrum of the Bølling warming by providing a plausible triggering mechanism for the onset of this event, traditionally considered as marking the termination of the last glacial period.

The Tahiti record

Here we report U–Th dating of coral samples collected from the Tahiti reef slope during the Integrated Ocean Drilling Program (IODP) Expedition 310, 'Tahiti Sea Level'²³. Tahiti is a far-field site located at a considerable distance from major former ice sheets and is characterized by slow and regular subsidence rates of ~0.25 mm yr⁻¹, as consistently assessed by several approaches. Considering a total range of 0.2–0.4 mm yr⁻¹ suggested by these approaches, the uncertainty on the assessment of the MWP-1A amplitude, arising from the correction of island subsidence during MWP-1A, is entirely negligible (see Supplementary Information). Previous reconstructions of the

¹CEREGE, UMR Aix-Marseille Université – CNRS – IRD – Collège de France, Technopole de l'Arbois, BP 80, 13545 Aix-en-Provence Cedex 4, France. ²Department of Earth Sciences, South Parks Road, Oxford OX1 3AN, UK. ³Atmosphere and Ocean Research Institute and Department of Earth and Planetary Science, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwashi, Chiba 277-8564, Japan. ⁴National Institute of Polar Research, Tachikawashi, Tokyo 190-8518, Japan. ⁵Institute of Biogeosciences, JAMSTEC, Yokosuka 237-0061, Japan.

deglacial sea-level rise were established from holes drilled onshore through the modern barrier-reef in front of Papeete harbour^{14,24}. The record was continuous from 13.9 kyr ago to present, but did not reach the critical MWP-1A period.

A specific target of Expedition 310 was the extension of the previous Tahiti sea-level record to cover earlier portions of the deglaciation. This was performed by offshore drilling of the Tahitian fore-reef slopes seaward of the present-day barrier reef (Fig. 1). These coring operations²³ recovered more than 400 m of post-glacial reef material, ranging from 122 to 40 metres below modern sea level (m.b.s.l.) in three distinct areas (Maraa, Faaa and Tiarei) around Tahiti (Fig. 1).

Our reconstruction of sea level relies on absolute U–Th dating of corals, belonging to corallgal (that is, coral and algal) assemblages indicative of a range of modern reef environments, from the shallow reef crest to the deepest reef slope. Eighty U–Th ages were determined on coral samples recovered from twenty-three holes drilled at fourteen different sites. These new data extend the Tahiti record to cover the last 16 kyr BP (Fig. 2), and provide a complete and detailed record of sea-level rise during this key period of the last deglaciation. In each hole, all of the ages are in stratigraphic order (Supplementary Fig. 2). However, even for the closely spaced holes, significant differences in recorded water depths may be observed (see, for example, the difference recorded between Site M0024 versus Site M0009 that may be up to ~10 m; Supplementary Fig. 2). The depth distribution observed for the various coral species analysed here is broadly consistent with their present-day biological zonation (Supplementary Fig. 4). The large number of holes drilled in the fore-reef slope, as well as their widespread distribution, ensured the recovery of the depth distribution of reef diversity and varying responses of reef development to sea-level rise. Our observations compare favourably with a reef accretion model²⁵, suggesting heterogeneous reef development induced by multiple factors including the following: spatially random (patchy) colonization; varying accretion patterns; and rugged topography of the pre-glacial surface that partially controlled the post-glacial reef initiation and growth following flooding. Our record, based on several contemporaneous cores, is therefore more representative than a record derived from a single drill hole which may provide a misleading impression of reef response to sea-level rise²⁵.

Sea-level rise during early deglaciation

The two oldest samples, dated at 15.74 ± 0.03 kyr BP and 16.09 ± 0.04 kyr BP, are robust branching *Pocillopora* collected at

the interface of the underlying Pleistocene unit in cores 24A-15R and 9B-15R. These samples belong to a shallow-water corallgal assemblage, <10 metres water depth (m.w.d.), and indicate a Relative Sea Level (RSL) of 117–107 m.b.s.l. during that time. This RSL estimate is strengthened by the presence of an encrusting *Montipora* collected at a subsidence-corrected depth of 114 m.b.s.l. in core 25B-11R. Dated at 15.31 ± 0.02 kyr BP, this sample is associated with vermetid gastropods that are indicative of a very shallow environment (<~5 m.w.d.)²⁶. From these observations, we may infer an RSL of 117–109 m.b.s.l. during the early part of the deglaciation at Tahiti (see Fig. 2).

Because of glacial isostatic adjustment (GIA), the RSL records from different sites cannot be compared directly, even in far-field regions²⁰. For the time window 14–20 kyr BP, GIA models produce an RSL that is lower at Tahiti than eustatic sea level^{20,27}, in contrast to other sites commonly used for the analysis of sea-level change (Barbados, Bonaparte Gulf and Huon Peninsula) where GIA effects lead to local sea level lying above the eustatic value. By taking this factor into account, our 117–109 m.b.s.l. RSL estimate at 16 kyr BP is therefore in good agreement with observations from the Sunda Shelf (Supplementary Fig. 8) for the same period⁸. RSL observations from Barbados and Bonaparte Gulf display a dense cluster of samples dated at about 18–19 kyr BP, which strongly constrains eustatic sea level to a depth less than 110 m in this interval²⁸. Therefore, a comparison with our data suggests that, during the early stage of deglaciation, after the MWP that occurred at 19 kyr BP^{9,10}, the eustatic sea level (ESL) remained stable or rose only slightly during the time span surrounding the Heinrich 1 event (probably no more than 5 m for ~3 kyr).

For the time window spanning 16.1–14.6 kyr BP, hole 24A (from the outer ridge at Tiarei) delineates the lower envelope of sea-level change. In this hole, corallgal assemblages are indicative of a very shallow environment and were able to keep pace with rising sea level during this period. The pre-MWP-1A RSL is well constrained by three coral samples collected at a subsidence-corrected depth of 105 m.b.s.l.: a massive *Montipora* sample dated at 14.65 ± 0.02 kyr BP in core 15A-37R from Maraa; and two robust branching *Pocillopora* samples dated at 14.58 ± 0.05 kyr BP and 14.61 ± 0.03 kyr BP in core 24A-10R from Tiarei (see Supplementary Information and Supplementary Fig. 3). These two last corals belong to a corallgal assemblage that typifies a shallow-water environment of less than 10 m.w.d. and are associated with vermetids that are indicative of shallow-water conditions (<~5 m.w.d.)²⁶. This places a conservative constraint of 105–100 m.b.s.l. on the pre-MWP-1A sea level at 14.65 kyr BP. A moderate

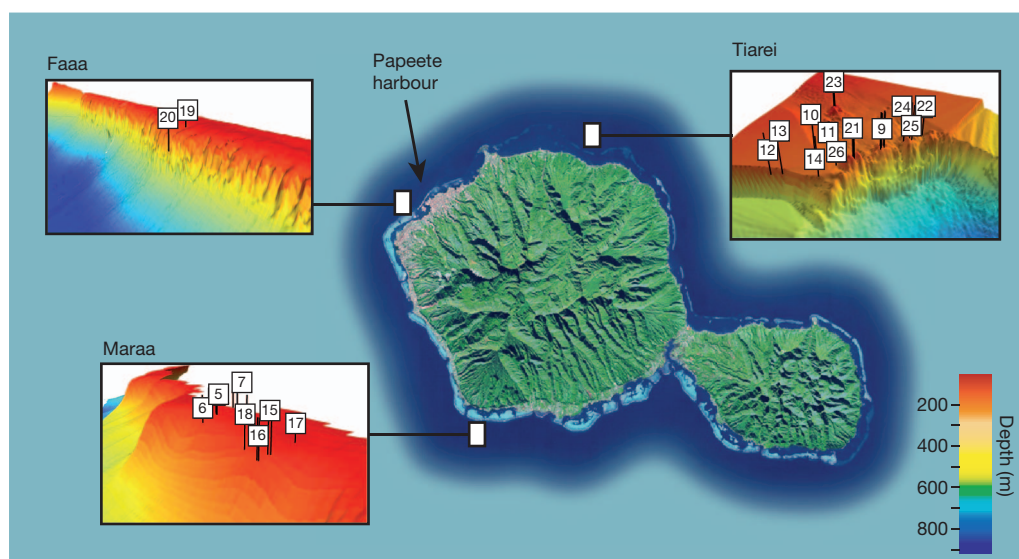


Figure 1 | A Landsat image of Tahiti island. Shown are the locations of the three areas (Tiarei, Maraa and Faaa) drilled during IODP Expedition 310, as well as Papeete harbour where onshore holes were drilled previously. A total of

37 boreholes were cored during IODP 310 at 22 different sites providing more than 400 m of post-glacial reef material²³. Insets show the bathymetry for each site, with the location of the different drilled holes.

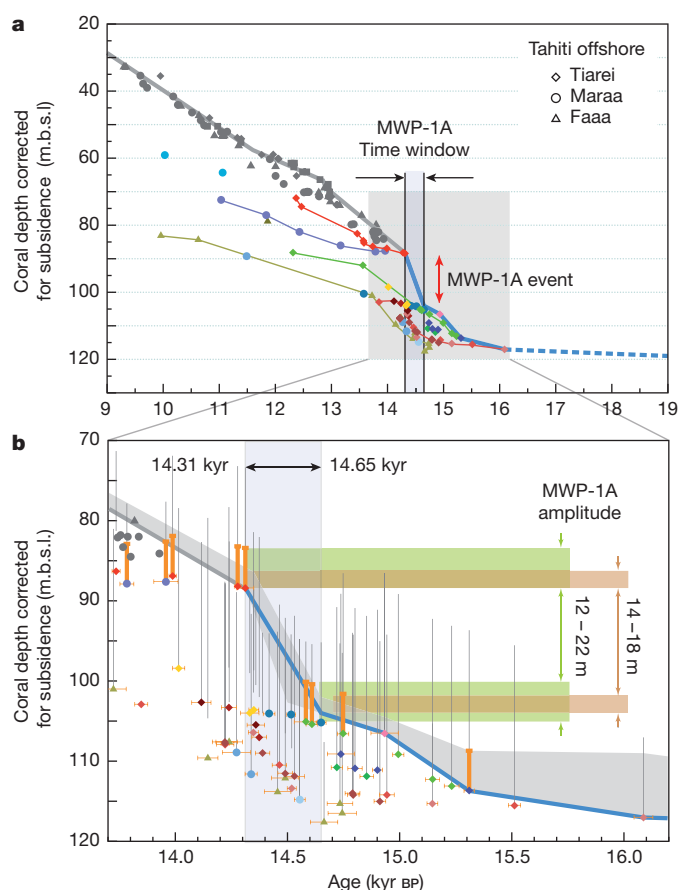


Figure 2 | The deglacial Tahiti sea-level curve. **a**, Sea level reconstructed from U–Th dated corals recovered in long holes drilled onshore and offshore Tahiti island. Coral depths are expressed in metres below present sea level (m.b.s.l.) and are corrected for a constant subsidence rate of 0.25 mm yr^{-1} (see Supplementary Information). Grey and coloured symbols show respectively coral samples collected in onshore holes^{14,24} and in offshore holes drilled during IODP Expedition 310. Red diamonds show key samples from the inner ridge of Tiarei (Site M0023). Thick blue line shows the lower estimate of the Tahiti RSL curve (see Supplementary Information); it extends the grey curve determined by linear fits of onshore sea-level data²⁴ and clearly indicates the occurrence of a rapid rise of the sea level (orange arrow) related to the MWP-1A event. The shaded time window and black arrows highlight the tight chronological constraints derived for MWP-1A from the Tahiti record. **b**, Magnified view of the MWP-1A time window. The vertical grey bars reported for each coral sample correspond to their optimal bathymetric habitat range inferred from the coral assemblage identification (see Supplementary Information) and thick orange bars indicate samples associated with vermetid gastropods that are indicative of a shallow environment ($0\text{--}5 \text{ m.w.d.}$). The shaded grey band illustrates our estimate of the most likely range of the Tahiti RSL over the last deglaciation. The ranges of uncertainty estimated from the bathymetric range of coral assemblages for the pre- and post-MWP-1A sea-level positions are illustrated by the horizontal green bands. The resulting extreme bounds for the MWP-1A amplitude (12 and 22 m) are also indicated (green bands and arrows). Several arguments given in the Supplementary Information suggest that these conservative estimates can be trimmed to 14 and 18 m (brown bands and arrows). Thick blue line and thick grey line are as in **a**.

sea-level rise of $4\text{--}14 \text{ m}$ is therefore inferred for the period from 16.1 to 14.65 kyr BP.

The earliest bound for the initiation of the MWP-1A jump of sea level is probably within the time range given by those three samples (14.58–14.65 kyr BP). Moreover, the two *Pocillopora* samples dated at 14.58 and 14.61 kyr BP could have already grown at a reasonable water depth (up to 5 m.w.d.). Thus, they may have already accommodated a part of the sea-level rise related to MWP-1A, implying that the inception of MWP-1A could have occurred somewhat earlier (see the upper

bound of the shaded grey area in Fig. 2; see also Supplementary Fig. 4). The maximum age for the onset of MWP-1A could thus be close to the oldest of these three corals, dated at 14.65 kyr BP. It must be emphasized that this only provides us with the uppermost limit for the onset of MWP-1A, and we cannot rule out that the jump may have started significantly later, as young as 14.5 kyr BP, as potentially marked by massive *Montipora* samples of core 15A-36R that characterize a shallow environment (see Supplementary Figs 3–5).

Occurrence of MWP-1A

The occurrence of MWP-1A is revealed by a major discontinuity in the upper envelope of the data points in the new Tahiti RSL record (Fig. 2). The next shallowest *in situ* samples in the sequence are two branching *Pocillopora* dated at $14.28 \pm 0.02 \text{ kyr BP}$ and $14.31 \pm 0.04 \text{ kyr BP}$ in cores 23B-12R and 23A-13R (see Supplementary Information and Supplementary Fig. 3). These coral samples, recovered at a subsidence-corrected depth of 88 m.b.s.l., are the first datable corals, showing clear evidence of an in-growth position, to colonize the pre-glacial substratum after the MWP-1A sea-level jump. These samples are critical, as they provide the most robust constraint on MWP-1A timing and clearly indicate that the sea-level jump was complete before 14.31 kyr BP. These data lie on the extension of the general trend depicted by onshore holes^{14,24} (Fig. 2) and highlight a regular, slow rate of sea-level rise after MWP-1A. These corals are associated with vermetids, thus indicating a very shallow environment ($<5 \text{ m.w.d.}$). We infer a conservative estimate of 88–83 m.b.s.l. for the post-MWP-1A sea level.

The MWP-1A event also coincides with a major change in reef development strategy, as illustrated by numerous samples dated in all drill holes collected on the outer edge of the fore-reef slopes. Before MWP-1A the reef kept pace with sea level, whereas a widespread deepening and backstepping occurred after MWP-1A. This change in reef response is coincident with changes in the coral assemblage composition, such as in Hole M0024A (see Supplementary Information and Supplementary Fig. 3), where shallow-water assemblages—dominated by robust branching *Pocillopora*, massive *Porites* and encrusting *Montipora*—change to branching *Porites* species, which typify an environment characterized by moderate energy and light intensity.

General features of reef geometry can be simulated with a two-dimensional growth model¹⁴. This model simulates the overall deepening of the reef sequence that follows occurrence of a rapid sea-level rise and clearly indicates that only holes drilled in the intermediate position between the outer ridge and the modern barrier reef are capable of capturing the sea-level position immediately following MWP-1A (see Supplementary Figs 9 and 10). This result probably explains the difficulty encountered by previous onshore or offshore drilling programmes (Tahiti or Barbados) to collect shallow-species coral samples that document precisely the end of MWP-1A. The IODP Mission Specific Platform overcame this difficulty by specifically targeting the reef structures located in intermediate position between the fore-reef slope and the present barrier reef, especially at the Tiarei site.

Amplitude and duration of MWP-1A at Tahiti

On the basis of the most conservative estimates deduced above for the pre- and post-MWP-1A sea level, we infer an amplitude of 17 m for the sea-level jump, with lowest and uppermost bounds of 12 and 22 m. Several arguments, discussed in detail in Supplementary Information (Supplementary Fig. 6), suggest that this range may reasonably be narrowed down to 14–18 m, with a median value of 16 m.

In view of the lower and upper limits of the MWP-1A chronozone (14.31 kyr BP and 14.65 kyr BP, respectively), the longest possible duration of the jump is ~ 350 years (Fig. 3). Considering the median value of 16 m for the local amplitude of MWP-1A at Tahiti, we infer an average RSL rate of $\sim 46 \pm 6 \text{ mm yr}^{-1}$ at Tahiti. However, owing to the age uncertainty associated with its inception and termination (see

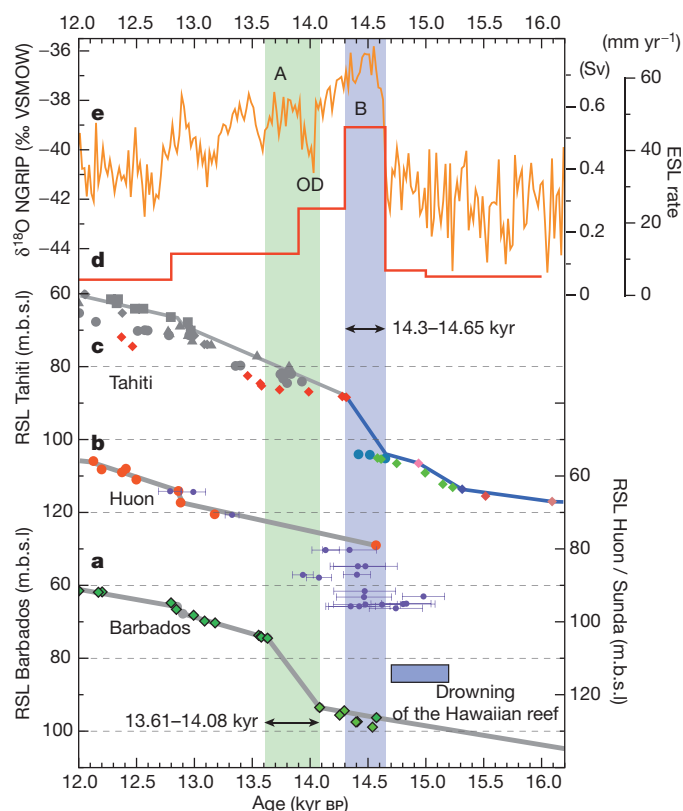


Figure 3 | Relative sea-level (RSL) records over the time window 16.5 to 12.0 kyr BP. **a**, Barbados RSL record based on U–Th dated corals (mainly *Acropora palmata*)^{19,29}. The shaded green vertical band highlights the MWP-1A time window inferred from the Barbados record^{19,29}. **b**, Pacific RSL records (right-hand vertical axis). Red circles, Huon Peninsula record^{31,32} (Papua New Guinea) based on U–Th dated corals. Purple points, Sunda Shelf record⁸ based on ¹⁴C-dated organic material found in sediment cores (recalibrated using IntCal09³⁰; plotted errors are 1σ). The blue rectangle indicates the drowning of a Hawaiian reef 14.7 kyr ago¹⁵. **c**, Tahiti RSL record based on U–Th dated corals collected in holes drilled onshore (grey symbols)^{14,24} and offshore (coloured symbols, this study). The shaded purple vertical band highlights the MWP-1A time window inferred from this study. **d**, Rate of glacial meltwater discharge (expressed in mm yr^{−1} and Sv, right-hand vertical axes) derived from the eustatic sea level curve determined by the GIA model (see Supplementary Information and Supplementary Fig. 11) adjusted to account for the newly obtained timing and magnitude of MWP-1A from Tahitian sea-level observations. **e**, δ¹⁸O record of the North Greenland Ice Core Project (NGRIP) core plotted on its most recent timescale³⁴; B, Bølling; OD, Older Dryas; A, Allerød. All depths have been corrected for subsidence (Tahiti) and uplift (all other sites) as described in ref. 24. For Tahiti and Barbados records, only samples that delineate the upper envelope are shown. Grey lines correspond to linear fits of sea-level data²⁴. Greenish and bluish shaded time windows correspond to MWP-1A chronozones inferred from the Barbados record and the Tahiti record (Fig. 2), respectively.

Supplementary Information), the MWP-1A duration could have been even shorter than this estimate. An extremely sharp meltwater outburst, of the order of a century or less, is thus possible, in which case the 46 mm yr^{−1} rate of sea-level rise must be considered as a minimum value.

Timing of MWP-1A

The new MWP-1A chronozone inferred from the extended Tahiti record (that is, 14.65–14.31 kyr BP or shorter, Fig. 3) does not overlap with that previously proposed on the basis of the Barbados record (14.08 ± 0.06 to 13.63 ± 0.03 kyr BP, using the most recent updated data set^{19,29}; see Supplementary Information for a full discussion of this issue).

Several other lines of evidence also suggest that MWP-1A was significantly older than suggested by the Barbados record and, ultimately, concurrent with the Bølling warming. Additional evidence comes from the Sunda Shelf sea-level record, derived from mangrove organic material collected from a shallow siliciclastic platform⁸. This record shows a very sharp sea-level rise dated at a conventional ¹⁴C age of 12.42 ± 0.06 kyr BP (1 s.d., *n* = 17; Supplementary Fig. 8) coinciding with the 500-year-long ¹⁴C plateau that encompasses the Bølling period. Using the IntCal09 calibration curve³⁰, the mean calendar age of the MWP-1A event recorded on the Sunda Shelf can be refined to 14.94–14.14 kyr cal. BP (2σ interval, see Supplementary Information for more details regarding this age calculation).

The revised MWP-1A timescale inferred from the new Tahiti record is also coherent with the recent extension of the Huon Peninsula record^{31,32}, where the oldest sample of the post-glacial reef sequence dated at 14.56 ± 0.05 kyr BP places an upper constraint on the end of MWP-1A (Fig. 3). Further indirect evidence is provided by the drowning of coral reefs offshore from Hawaii, which occurred at 14.7 kyr BP and has been proposed to be caused by a dramatic increase in sea level related to MWP-1A¹⁵.

These records are consistent enough to revise the onset of MWP-1A so it is 500 years earlier than the date inferred from the Barbados data. Within this revised timeframe, MWP-1A can no longer be advocated as the trigger for the Older Dryas cooling event that terminated the Bølling period, as proposed previously^{14,16,33}. Instead, MWP-1A coincided with the inception of the Bølling period (Fig. 3), which has been independently constrained by the GICC 05 Greenland ice core chronology at 14.640 kyr BP (with a maximum counting error of 0.186 kyr)³⁴. The Tahiti record is thus compatible with the idea of a temporal relationship between MWP-1A and Bølling warming. This hypothesis is further substantiated by the concurrent occurrence of rapid flooding on shelf margins and an increase in sea surface temperature in the South China Sea at the Bølling transition²¹.

Source of MWP-1A

Because they account for more than 80% of total sea-level rise during the last deglaciation, NHIS, and especially the Laurentide Ice Sheet (LIS), have commonly been considered as the sole sources for MWP-1A^{5,35}. But arguments for such an LIS source faced serious objections, and led to the proposal¹⁷ of an alternative scenario in which a significant fraction of the melt water came from Antarctica.

Direct evidence in favour of a Northern or Southern Hemisphere source remains equivocal. Most robust arguments supporting an Antarctic contribution were provided by GIA models^{20,36,37}. Fingerprinting model experiments demonstrated that comparison of the size of the MWP-1A sea-level rise observed at several sites could provide helpful information about the source(s) of melting ice³⁶. Predictions provided in ref. 36 showed that, when melting ice originated exclusively from the LIS, the amplitude of MWP-1A predicted for Barbados should be significantly lower than for far-field sites. This scenario predicted the greatest difference in amplitude between Barbados and Tahiti, with a sea-level rise at Tahiti almost twice that at Barbados³⁶.

The amplitude of MWP-1A that we assess at Tahiti (16 m) is comparable to that observed at Sunda (~16 m)⁸. At Barbados, the amplitude of the jump must be reassessed on the basis of the re-evaluation of the MWP-1A chronozone (Supplementary Fig. 7). By extrapolating the linear trend defined by hole 12 (Supplementary Fig. 7), we roughly estimate a ~15 m amplitude of sea-level rise at Barbados. The amplitudes of MWP-1A recorded at these three far-to-intermediate-field sites are thus approximately the same. Following the predictions of ref. 36, our results seem to preclude a sole LIS contribution to MWP-1A and confirm the preliminary conclusions³⁶ based solely on the Sunda and Barbados records. On this basis, the Barents and Fennoscandian Ice Sheets can also be considered as possible candidates for the freshwater source (see figure 2 in ref. 36), but there are

several counterarguments to these ice sheets as the major sources of fresh water¹⁷. All other scenarios that provide equal amplitudes of MWP-1A sea-level rise require a significant Antarctic contribution.

These arguments in favour of a contribution from the AIS were reinforced by GIA predictions²⁰. Those predictions showed that the optimal deglacial scenario to fit RSL observations at Barbados, Tahiti, Huon Peninsula and Sunda Shelf during late glacial time required a MWP-1A with a total amplitude of 23 m, which included an AIS contribution of 15 m with a total NHIS contribution of 8 m (6 m from the LIS).

Using a realistic GIA model (see Supplementary Fig. 11 and Supplementary Information), which uses the Earth model proposed in ref. 20, we performed a new set of simulations that agree well with the conclusion of Bassett *et al.*²⁰, pointing towards a substantial contribution from the AIS. It is difficult at this stage, however, to conclusively determine the relative contributions of NHIS and the AIS to MWP-1A because these approaches (fingerprinting and more general GIA modelling) are hampered by uncertainties surrounding the MWP-1A-induced relative sea-level amplitude, especially at the intermediate-field site of Barbados. Following previous studies^{27,36}, which conclude that the MWP-1A amplitude recorded at Tahiti is amplified by 10–30% with respect to its eustatic amplitude, our results are consistent with a eustatic MWP-1A rise of roughly ~ 14 m during the time window 14.65–14.3 kyr BP, leading to a rate of eustatic sea level rise of 40 mm yr^{-1} . Note that this value is significantly lower than the 20–25 m of eustatic rise often reported in the literature^{20,36}. Considering the growing body of evidence^{20,36,37} that suggests that a substantial fraction of MWP-1A originated from Antarctica, it is probable that the AIS contributed at least half of the ~ 14 m eustatic sea-level rise observed during this event. It is worth noting that this estimate of the Antarctic contribution allows us to balance the freshwater budget required for MWP-1A, taking into account NHIS contributions that have been independently assessed to be between 5 and 10 m of sea-level equivalent ice volume^{38,39}. Recent estimates of AIS contribution to the last deglaciation indicate that its contribution was < 20 m and perhaps lower than 10–15 m (refs 40–42), implying that a significant, if not the major, part of the AIS contribution to the last deglaciation occurred during MWP-1A.

Implications of the revisited MWP-1A history

The IODP Expedition 310 provides significantly improved constraints on the timing of MWP-1A, demonstrating that MWP-1A ended before 14.3 kyr BP and that it started after 14.65 kyr BP. This makes MWP-1A coeval with the Bølling warming, suggesting a temporal, and probably causal, relationship between these two prominent deglacial features. Owing to the dating uncertainty of the Bølling inception in the Greenland ice record (14.642 kyr BP with a maximum counting error of 186 years; ref. 34), it remains difficult to unravel the phasing and causal mechanisms linking—through specific atmospheric and oceanic responses—the resumption of the AMOC during the Bølling warming²² and massive meltwater discharges in both hemispheres. Two end-member scenarios that warrant further investigation can be put forward, however:

The first scenario is that proposed in ref. 12, based on GCM simulations showing that a rapid freshwater discharge originating from the AIS could have led to an intensification of the AMOC. The associated northward ocean heat flux would trigger the Bølling warming in the Northern Hemisphere and a rapid melting of the LIS. But subsequent studies (for example, ref. 43 and references therein) that have tested the scenario of ref. 12 showed that the meltwater discharge may have led to competing mechanisms, enhancing or weakening the AMOC, which collectively lead to a subdued climatic response in the Northern Hemisphere⁴³.

In the second scenario, the phasing of events is reversed, with an initial AMOC increase and associated northward ocean heat transport causing the Bølling warming, which led to rapid melting of NHIS,

in particular the LIS. The resulting sea-level rise drove in turn a dramatic collapse of the AIS. Indeed, the WAIS was partly marine-based during the LGM and thus probably sensitive to the break-up and loss of buttressing ice shelves. In any case, most of the WAIS is characterized by unstable conditions, with bedrock below sea level and slopes downward from the margins towards the interior¹³.

In fact, these two scenarios are not mutually exclusive and could have acted in concert during the MWP-1A chronozone, reinforcing each other. They are both compatible with our sea-level and source fingerprinting study, which implies that meltwater injections forming the MWP-1A event originated from ice sheets in both Antarctica and the Northern Hemisphere, including the LIS. In principle, meltwater injection into the North Atlantic could have counteracted the AMOC increase, but the strength of this negative feedback depends on the exact location and mode of meltwater release. Several studies suggested that LIS meltwater was funnelled through the Mississippi drainage system, before being released in the Gulf of Mexico as a hyperpycnal flow^{38,44}, with a negligible impact on the AMOC^{39,45,46}.

The two scenarios have similar ingredients but differ in their ultimate trigger, AIS collapse or AMOC increase. These abrupt events could be linked to threshold responses to the gradual warming of the Southern Hemisphere that occurred under external forcings (orbital and greenhouse-gas changes) during the early part of the deglaciation^{47,48}.

Much research remains to be done to document the precise sequence of events during the MWP-1A chronozone. This will come from coring coral reefs at other sites (for example, Barbados and the Seychelles²⁷), from study of open-ocean sediments in the vicinity of former ice sheets, and from modelling work to simulate the complex interplay between ice sheets, ocean and atmosphere. Whatever the causes that led to the MWP-1A event and the Bølling warming, and despite the fact that the total eustatic magnitude of this event is reduced compared to previous estimates, our results prove the existence of a dramatic collapse of past ice sheets at a eustatic rate exceeding 40 mm yr^{-1} , with a substantial contribution from Antarctica. We note that this rate is at least four times as large as the average rate of deglacial sea-level rise of $\sim 10 \text{ mm yr}^{-1}$; see ref. 24 and Supplementary Information. Understanding this singular event will shed light on the dynamical behaviour of large ice sheets in response to external forcing or internal perturbation of the climate system. This topic is crucial in the context of the present warming, as modern ice sheets have been shown to be contributing directly to the recent acceleration in sea-level rise^{1,2}.

METHODS SUMMARY

Before U–Th dating, rigorous mineralogical and isotopic screening criteria were applied to discard coral samples that suffered any post-mortem diagenetic alteration of their aragonite skeleton. In particular, using X-ray diffraction³⁰, we made an effort to improve the detection and quantification of a very small amount of secondary calcite. Coral samples showing a calcite content of more than 1% were discarded. Most of the U–Th analyses were performed using a VG-54 thermionization mass spectrometer equipped with a 30-cm electrostatic analyser and a pulse-counting Daly detector at CEREGE (see Supplementary Information for data and analytical issues). The initial $(^{234}\text{U}/^{238}\text{U})_0$ values calculated for post-glacial samples yielded a mean value of 1.1458 ± 0.0020 (2σ), falling within the most recent determinations of modern sea water and corals⁴⁹. Additionally, for corals of the same age, $(^{234}\text{U}/^{238}\text{U})_0$ values were highly consistent (that is, within an analytical uncertainty determined for the entire course of the study of 0.8%, 2σ), and within the larger range adopted³⁰ as an isotopic screening criterion in the interval 0–17 kyr BP ($(^{234}\text{U}/^{238}\text{U})_0 = 1.1452 \pm 0.0048$, 2σ). The clustering of $(^{234}\text{U}/^{238}\text{U})_0$ values determined in this study substantially narrows the uncertainty for the evolution of the seawater value through time compared to previous data sets (Vanuatu, Papua New Guinea and Barbados) that have encompassed the last deglaciation, highlighting the outstanding quality of the coral samples recovered in Tahiti offshore holes. Complementary and duplicated analyses were also performed by Multi-Collector Inductively Coupled Mass Spectrometry⁵⁰ and show a general good agreement within measurement uncertainties.

Received 9 July 2011; accepted 26 January 2012.

1. Milne, G. A., Gehrels, W. R., Hughes, C. W. & Tamisiea, M. E. Identifying the causes of sea-level change. *Nature Geosci.* **2**, 471–478 (2009).
2. Pfeffer, W. T., Harper, J. T. & O'Neel, S. Kinematic constraints on glacier contributions to 21st-century sea-level rise. *Science* **321**, 1340–1343 (2008).
3. Mehl, G. A. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 747–845 (Cambridge Univ. Press, 2007).
4. Alley, R. B., Clark, P. U., Huybrechts, P. & Joughin, I. Ice-sheet and sea-level changes. *Science* **310**, 456–460 (2005).
5. Fairbanks, R. G. A 17,000-year glacio-eustatic sea level record; influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature* **342**, 637–642 (1989).
6. Bard, E., Hamelin, B. & Fairbanks, R. G. U-Th ages obtained by mass spectrometry in corals from Barbados: sea level during the past 130,000 years. *Nature* **346**, 456–458 (1990).
7. Blanchon, P. & Shaw, J. Reef drowning during the last deglaciation: evidence for catastrophic sea-level rise and ice-sheet collapse. *Geology* **23**, 4–8 (1995).
8. Hanebuth, T., Stattegger, K. & Grootes, P. M. Rapid flooding of the Sunda Shelf: a late-glacial sea-level record. *Science* **288**, 1033–1035 (2000).
9. Yokoyama, Y., Lambeck, K., De Deckker, P., Johnston, P. & Fifield, L. K. Timing of the Last Glacial Maximum from observed sea-level minima. *Nature* **406**, 713–716 (2000).
10. Clark, P. U., McCabe, A. M., Mix, A. C. & Weaver, A. J. Rapid rise of sea level 19,000 years ago and its global implications. *Science* **304**, 1141–1144 (2004).
11. Manabe, S. & Stouffer, R. J. Simulation of abrupt climate change induced by freshwater input to the North Atlantic Ocean. *Nature* **378**, 165–167 (1995).
12. Weaver, A. J., Saenko, O. A., Clark, P. U. & Mitrovica, J. X. Meltwater pulse 1A from Antarctica as a trigger of the Bølling-Allerød warm interval. *Science* **299**, 1709–1713 (2003).
13. Bamber, J. L., Riva, R. E. M., Vermeersen, B. L. A. & LeBrocq, A. M. Reassessment of the potential sea-level rise from a collapse of the West Antarctic Ice Sheet. *Science* **324**, 901–903 (2009).
14. Bard, E. *et al.* Deglacial sea-level record from Tahiti corals and the timing of global meltwater discharge. *Nature* **382**, 241–244 (1996).
15. Webster, J. M. *et al.* Drowning of the –150 m reef off Hawaii: a casualty of global meltwater pulse 1A? *Geology* **32**, 249–252 (2004).
16. Stanford, J. D. *et al.* Timing of meltwater pulse 1a and climate responses to meltwater injections. *Paleoceanography* **21**, PA4103, <http://dx.doi.org/10.1029/2006PA001340> (2006).
17. Clark, P. U. *et al.* Origin of the first global meltwater pulse following the last glacial maximum. *Paleoceanography* **11**, 563–577 (1996).
18. Peltier, W. R. On the hemispheric origins of meltwater pulse 1a. *Quat. Sci. Rev.* **24**, 1655–1671 (2005).
19. Peltier, W. R. & Fairbanks, R. G. Global glacial ice volume and Last Glacial Maximum duration from an extended Barbados sea level record. *Quat. Sci. Rev.* **25**, 3322–3337 (2006).
20. Bassett, S. E., Milne, G. A., Mitrovica, J. X. & Clark, P. U. Ice sheet and solid earth influences on far-field sea-level histories. *Science* **309**, 925–928 (2005).
21. Kienast, M., Hanebuth, T. J. J., Pelejero, C. & Steinke, S. Synchronicity of meltwater pulse 1a and the Bølling warming: new evidence from the South China Sea. *Geology* **31**, 67–70 (2003).
22. McManus, J. F., Francois, R., Gherardi, J. M., Keigwin, L. D. & Brown-Leger, S. Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature* **428**, 834–837 (2004).
23. Camoin, G., Iryu, Y., McInroy, D., & Expedition 310 Scientists. Expedition 310. *Proc. IODP Vol. 310 Expedition Reports* <http://dx.doi.org/10.2204/iodp.proc.310.2007> (2007).
24. Bard, E., Hamelin, B. & Delanghe-Sabatier, D. Deglacial meltwater pulse 1B and Younger Dryas sea levels revisited with boreholes at Tahiti. *Science* **327**, 1235–1237 (2010).
25. Blanchon, P. & Blakeway, D. Are catch-up reefs an artefact of coring? *Sedimentology* **50**, 1271–1282 (2003).
26. Cabioch, G., Montaggioni, L. F., Faure, G. & Ribaud-Laurenti, A. Reef coralgal assemblages as recorders of paleobathymetry and sea level changes in the Indo-Pacific province. *Quat. Sci. Rev.* **18**, 1681–1695 (1999).
27. Milne, G. A. & Mitrovica, J. X. Searching for eustasy in deglacial sea-level histories. *Quat. Sci. Rev.* **27**, 2292–2302 (2008).
28. Peltier, W. R. On eustatic sea level history: Last Glacial Maximum to Holocene. *Quat. Sci. Rev.* **21**, 377–396 (2002).
29. Fairbanks, R. G. *et al.* Radiocarbon calibration curve spanning 0 to 50,000 years BP based on paired $^{230}\text{Th}/^{234}\text{U}$ and ^{14}C dates on pristine corals. *Quat. Sci. Rev.* **24**, 1781–1796 (2005).
30. Reimer, P. J. *et al.* Intcal09 and Marine09 radiocarbon age calibration curves, 0–50,000 years cal BP. *Radiocarbon* **51**, 1111–1150 (2009).
31. Edwards, R. L. *et al.* A large drop in atmospheric $^{14}\text{C}/^{12}\text{C}$ and reduced melting in the Younger Dryas, documented with ^{230}Th ages of corals. *Science* **260**, 962–968 (1993).
32. Cutler, K. B. *et al.* Rapid sea-level fall and deep-ocean temperature change since the last interglacial period. *Earth Planet. Sci. Lett.* **206**, 253–271 (2003).
33. Liu, J. P. & Milliman, J. D. Reconsidering meltwater pulses 1A and 1B: global impacts of rapid sea level rise. *J. Ocean Univ. China* **3**, 183–190 (2004).
34. Rasmussen, T. L. *et al.* A new Greenland ice core chronology for the last glacial termination. *J. Geophys. Res.* **111**, D06102, <http://dx.doi.org/10.1029/2005JD006079> (2006).
35. Peltier, W. R. Ice-age paleotopography. *Science* **265**, 195–201 (1994).
36. Clark, P. U., Mitrovica, J. X., Milne, G. A. & Tamisiea, M. E. Sea-level fingerprinting as a direct test for the source of global meltwater pulse 1A. *Science* **295**, 2438–2441 (2002).
37. Bassett, S. E., Milne, G. A., Bentley, M. J. & Huybrechts, P. Modelling Antarctic sea-level data to explore the possibility of a dominant Antarctic contribution to meltwater pulse 1A. *Quat. Sci. Rev.* **26**, 2113–2127 (2007).
38. Carlson, A. E. Geochemical constraints on the Laurentide Ice Sheet contribution to Meltwater Pulse 1A. *Quat. Sci. Rev.* **28**, 1625–1630 (2009).
39. Tarasov, L. & Peltier, W. R. A calibrated deglacial drainage chronology for the North American continent: evidence of an Arctic trigger for the Younger Dryas. *Quat. Sci. Rev.* **25**, 659–688 (2006).
40. Clark, P. U. & Mix, A. C. Ice sheets and sea level of the Last Glacial Maximum. *Quat. Sci. Rev.* **21**, 1–7 (2002).
41. Denton, G. H. & Hughes, T. J. Reconstructing the Antarctic Ice Sheet at the Last Glacial Maximum. *Quat. Sci. Rev.* **21**, 193–202 (2002).
42. Bentley, M. J. The Antarctic palaeo record and its role in improving predictions of future Antarctic Ice Sheet change. *J. Quat. Sci.* **25**, 5–18 (2010).
43. Swingedouw, D., Fichefet, T., Goosse, H. & Loutre, M. F. Impact of transient freshwater releases in the Southern Ocean on the AMOC and climate. *Clim. Dyn.* **33**, 365–381 (2009).
44. Aharon, P. Entrainment of meltwaters in hyperpycnal flows during deglaciation superstorms in the Gulf of Mexico. *Earth Planet. Sci. Lett.* **241**, 260–270 (2006).
45. Roche, D. M., Renssen, H., Weber, S. L. & Goosse, H. Could meltwater pulses have been sneaked unnoticed into the deep ocean during the last glacial? *Geophys. Res. Lett.* **34**, L24708, <http://dx.doi.org/10.1029/2007GL032064> (2007).
46. Tarasov, L. & Peltier, W. R. Arctic freshwater forcing of the Younger Dryas cold reversal. *Nature* **435**, 662–665 (2005).
47. Knorr, G. & Lohmann, G. Southern Ocean origin for the resumption of Atlantic thermohaline circulation during deglaciation. *Nature* **424**, 532–536 (2003).
48. Knorr, G. & Lohmann, G. Rapid transitions in the Atlantic thermohaline circulation triggered by global warming and meltwater during the last deglaciation. *Geochim. Geophys. Res.* **8**, Q12006, <http://dx.doi.org/10.1029/2007GC001604> (2007).
49. Andersen, M. B. *et al.* The tracing of riverine U in Arctic seawater with very precise $^{234}\text{U}/^{238}\text{U}$ measurements. *Earth Planet. Sci. Lett.* **259**, 171–185 (2007).
50. Thomas, A. L. *et al.* Penultimate deglacial sea-level timing from uranium/thorium dating of Tahitian corals. *Science* **324**, 1186–1189 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the IODP and ECORD (European Consortium for Ocean Research Drilling) for drilling offshore from Tahiti, and the Bremen Core Repository members for organizing the onshore sampling party. We dedicate this article to the memory of G. Cabioch, who died at the end of 2011: a reef geology expert, he was a member of the IODP expedition to Tahiti. The CEREGE group thanks W. Barthelmy for maintaining mass spectrometers; D. Borschneck for help with X-ray diffraction analyses; and P. Dussouillez for help with maps. Palaeoclimate work at CEREGE is supported by the Comer Science and Education Foundation, the European Science Foundation (EuroMARC), the European Community (Project Past4Future), the Collège de France and the IRD (Institut de Recherche pour le Développement). The Oxford University team is supported by UK Natural Environment Research Council grant NE/D001250/1 and the Comer Science and Education Foundation. The University of Tokyo group is partly supported by JSPS (NEXT program GR031).

Author Contributions G.C., E.B. and B.H. were Principal Investigators for ODP proposal 519 designing this study. G.C. was co-chief scientist of IODP Expedition 310. P.D. and A.L.T. participated in the IODP sampling party. N.D., P.D. and A.L.T. performed U-Th dating of coral samples; N.D. performed X-ray diffraction analyses and reef growth modelling simulations; J.O. and Y.Y. performed geophysical modelling simulations; P.D. wrote the manuscript in collaboration with E.B. and B.H. The paper was refined by contributions from N.D., A.L.T., G.M.H. and G.C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.D. (deschamps@cerege.fr).

A new hominin foot from Ethiopia shows multiple Pliocene bipedal adaptations

Yohannes Haile-Selassie^{1,2}, Beverly Z. Saylor², Alan Deino³, Naomi E. Levin⁴, Mulugeta Alene⁵ & Bruce M. Latimer²

A newly discovered partial hominin foot skeleton from eastern Africa indicates the presence of more than one hominin locomotor adaptation at the beginning of the Late Pliocene epoch. Here we show that new pedal elements, dated to about 3.4 million years ago, belong to a species that does not match the contemporaneous *Australopithecus afarensis* in its morphology and inferred locomotor adaptations, but instead are more similar to the earlier *Ardipithecus ramidus* in possessing an opposable great toe. This not only indicates the presence of more than one hominin species at the beginning of the Late Pliocene of eastern Africa, but also indicates the persistence of a species with *Ar. ramidus*-like locomotor adaptation into the Late Pliocene.

Woranso-Mille is a relatively new palaeontological site located in the central Afar region of Ethiopia¹. The fossiliferous horizons identified at the site range in age from approximately 3.2 to 3.8 million years (Myr) ago. More than 54,000 fossil specimens sampling diverse mammalian taxa have been collected thus far (Supplementary Information). Geological and palaeontological work in the past five years has concentrated on sediments radiometrically dated to between 3.57 ± 0.014 and 3.8 ± 0.18 Myr ago². These sediments have yielded numerous early hominin remains, including a partial skeleton of *Au. afarensis*^{3–5}. Slightly younger deposits have subsequently yielded hominin fossils including a well-preserved, ~3.4-Myr-old partial foot skeleton (BRT-VP-2/73). The detailed geological context, dating and palaeoenvironment of BRT-VP-2/73 are presented in the Supplementary Information.

The hominin forefoot (metatarsals and phalanges) is characteristically under-represented in the fossil record as a consequence of its fragility in the face of predators and taphonomic processes. Previously described hominin pedal fossils^{6–12} have not included associated and well-preserved metatarsals and phalanges. Here we describe a partial hominin forefoot (BRT-VP-2/73) recovered from Burtele locality 2 (BRT-VP-2), one of the vertebrate localities of the Woranso-Mille study area (see Fig. 1). This partial pedal skeleton is unique in providing important evidence bearing on the functional morphology and proportions of several early hominin foot elements. It also presents the opportunity to draw morphological and functional comparisons between earlier (*Ar. ramidus*, ~4.4 Myr ago) and contemporaneous (*Au. afarensis*, ~2.9–3.6 Myr ago) hominins, and test whether there was diversity in hominin bipedalism in the earlier phases of hominin evolutionary history¹³.

BRT-VP-2/73 consists of eight mostly intact bony elements of a right foot: complete first, second, fourth metatarsals; head of third metatarsal; three proximal phalanges (rays 1, 2 and 4); and one middle phalanx (ray 2) (Fig. 2a–f and Table 1). Detailed comparative descriptions are provided in Supplementary Information. The lack of anatomical redundancy, spatial distribution, individual age status, morphological compatibility and preservation of the specimens indicate that they are from a single foot.

BRT-VP-2/73 clearly differs from cercopithecids by its dorsoplantarly tall hallux base relative to the bone's length (Fig. 3a) and also relative to the height of the second metatarsal base (Fig. 3b), in addition to a number

of other metatarsal ratios (Fig. 4, see Supplementary Information for discussions). Principal components analysis (PCA; correlation matrix, varimax rotation with Kaiser normalization) was conducted on 11 metatarsal ratios (Supplementary Table 1). Although some metatarsal length proportions of BRT-VP-2/73 are more similar to those of cercopithecids (for example, MT2 length < MT4 length) than those

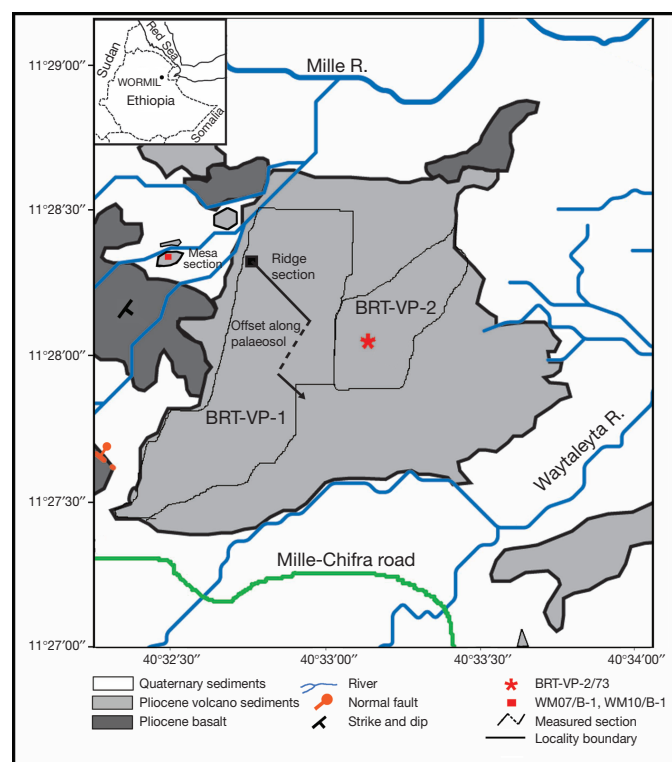


Figure 1 | Location map of the Burtele (BRT) vertebrate localities (BRT-VP-1 and BRT-VP-2) in the Woranso-Mille study area. The path of the measured section through the sandstone ridges and the location of the mesa section with the dated Burtele tuff are shown. The measured basalt section is off the map. The study area is located about 30 miles north of Hadar and Gona.

¹The Cleveland Museum of Natural History, Cleveland, Ohio 44106, USA. ²Case Western Reserve University, Cleveland, Ohio 44106, USA. ³Berkeley Geochronology Center, Berkeley, California 94720, USA. ⁴Johns Hopkins University, Baltimore, Maryland 21218, USA. ⁵Addis Ababa University, PO Box 1176 Addis Ababa, Ethiopia.



Figure 2 | Pedal elements of BRT-VP-2/73. **a**, Dorsal view of all elements of the specimen. **b**, Dorsal, plantar, lateral, medial, distal and proximal views of the first metatarsal. **c**, Dorsal, lateral, medial, proximal and distal views of the second metatarsal. **d**, Dorsal, lateral, plantar, distal and proximal views of the

hallux proximal phalanx. **e**, Lateral views of the second and fourth proximal phalanges, and the second intermediate phalanx. **f**, Dorsal, plantar and lateral views of the fourth metatarsal. All views are from left to right.

of apes or humans, the results of the PCA clearly distinguish BRT-VP-2/73 from Old World monkeys and show that it falls in the cluster formed by anatomically modern humans and gorillas (Fig. 4 and Supplementary Information).

The proximal phalanges of BRT-VP-2/73 show the pronounced dorsal canting associated with substantial doming of its metatarsal heads similar to the condition seen in humans and early hominins such as *Ar. ramidus* and *Au. afarensis*. BRT-VP-2/73 differs from chimpanzees by lacking long and curved metatarsal shafts (Supplementary Fig. 1) and phalanges, and in having a larger degree of dorsiflexion at the lateral metatarsophalangeal joints. It also differs from African apes by the degree of torsion of its hallux head (Supplementary Fig. 2) and doming of its second and fourth metatarsal heads. BRT-VP-2/73 is similar to *Ar. ramidus* in showing a mosaic of derived hominin pedal characteristics associated with obligate bipedality and other features associated with arboreality. For example, it resembles *Ar. ramidus* in combining an abducted hallux and medially directed torsion of the second metatarsal. However, its attribution to this species would be premature particularly in the absence of associated craniodental elements.

Comparative description

The hallux is represented by a complete, well-preserved, right first metatarsal (BRT-VP-2/73c) and its associated proximal phalanx (BRT-VP-2/73g; Fig. 2a, b, d). The articular base of the metatarsal is tall, deeply concave, and it exhibits the sigmoidal configuration seen in extant African apes and in *Ar. ramidus*¹¹. There is a low ridge running obliquely across the proximal articular surface from its medial dorsoplantar midpoint to the attachment area of the fibularis longus. A similar feature sometimes occurs in *Gorilla* hallux metatarsals. This subdued ridge is not like that described in the proximal metatarsal base from Hadar, Ethiopia (A.L. 333-54) wherein a distinct elevation nearly horizontally bisects the articular base into two semicircular facets¹⁴. The BRT-VP-2/73c base is notably tall relative to the bone's length, exceeding the ranges in chimpanzees and Old World monkeys, but within the ranges of gorillas and anatomically modern humans for this ratio (Fig. 3a).

The BRT-VP-2/73c metatarsal head does not conform to the 'typical' *Australopithecus* pattern in lacking the dramatic dorsal doming that characterizes this genus^{6,14} (for example, A.L. 333-115a and A.L. 333-21).

Table 1 | Linear measurements of the pedal elements of BRT-VP-2/73

Specimen no.	Element	M1 (mm)	M2 (mm)	M3 (mm)	M4 (mm)	M5 (mm)	M6 (mm)	M7 (mm)	M8 (°)
BRT-VP-2/73a	R. MT4	68.7	12.7	13.3	10.5	12.1	5.4	9.2	26–27†
BRT-VP-2/73b	R. MT2	66.9	12.8	14.2	9.8	11.2	6.05	7.35	23‡
BRT-VP-2/73c	R. MT1	50.3	14.6	22.7	16.7	14.5	9.05	8.95	–
BRT-VP-2/73d	R. prox. PHX 4	28.74	10.25	8.6	7.9	5.4	5.32	5.16	–
BRT-VP-2/73e	R. prox. PHX 2	29.7	10.9	9.6	7.95	5.3	6.35	6.02	–
BRT-VP-2/73f	R. MT3 head	15.1*	–	–	8.6*	13.2*	–	–	–
BRT-VP-2/73g	R. prox. PHX 1	25.23	13.1	9.73	12.24	6.5	8.45	6.06	–
BRT-VP-2/73h	R. Int. PHX 2	18.5	9.26	7.63	7.3	4.4	5.1	3.85	–

M1, maximum length; M2, proximal articular joint mediolateral; M3, proximal articular joint dorsoplantar; M4, distal articular joint mediolateral; M5, distal articular joint dorsoplantar; M6, midshaft mediolateral; M7, midshaft dorsoplantar; M8, distal head torsion.
* Preserved dimension.
† Lateral torsion in degrees.
‡ Medial torsion in degrees.

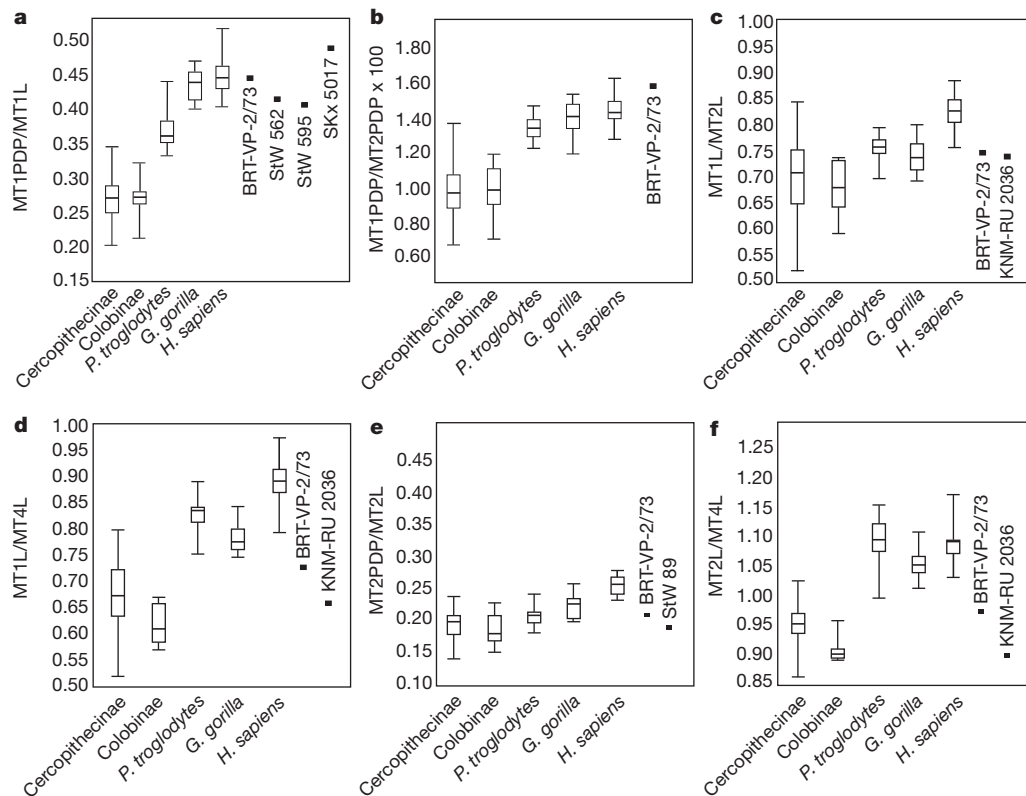


Figure 3 | Box-and-whisker plots of pedal element comparative ratios in cercopithecines, colobines, chimpanzees, gorillas, humans and fossil specimens. (See Supplementary Table 6 for taxonomic composition.) Whisker lines indicate maximum and minimum values. **a**, Base height (PDP) of the first metatarsal to its length (L). **b**, Base height of the first metatarsal to base height of the second metatarsal. **c**, Hallucal length to the second metatarsal length. **d**, Hallucal length to the fourth metatarsal length. **e**, Base height of the second metatarsal to its length. **f**, Length of the second metatarsal to the length of the fourth metatarsal. Measurements of the South African and Miocene hominoids were taken from refs 10 and 17, respectively.

Its dorsoproximal articular margin is continuous and it does not exhibit the 'nonsubchondral isthmus' described in *Ar. ramidus*¹¹.

A simple ratio comparing the length of the first metatarsal to the lengths of the second and fourth metatarsals demonstrates that the

hallucal segment is relatively short, falling within the ranges of the African apes (Fig. 3c, d) and outside the range for anatomically modern humans (Supplementary Table 2). However, its tall hallucal base, relative to the shorter bases of the associated metatarsals, indicates that the BRT foot had a transverse arch more developed than in apes and falls in this ratio at the higher range for anatomically modern humans (Fig. 3a).

The hallucal proximal phalanx is essentially complete and, when combined with its associated metatarsal, further confirms that the hallucal ray is relatively short. A ratio formed between the combined lengths of the first metatarsal and its associated proximal phalanx (MT1 + PP1) and the same elements from the second ray (MT2 + PP2) demonstrates that anatomically modern humans with their elongated halluces are notably distinct. BRT-VP-2/73c falls within the ranges of apes and monkeys, indicating that the foot had a relatively short, abductable great toe (Supplementary Fig. 3). This ratio also confirms that the BRT-VP-2/73 hallucal ray was not used during a human-like toe-off in the terminal phase of the gait cycle. However, the degree of its proximal joint canting (97°) is lower than in the second ray (100°), which is a condition seen in humans, whereas the opposite is the case in chimpanzees¹⁵ (Supplementary Fig. 4a).

The second ray is represented by a metatarsal (BRT-VP-2/73b), a proximal phalanx (BRT-VP-2/73e) and an intermediate phalanx (BRT-VP-2/73h). BRT-VP-2/73b is a well-preserved second metatarsal. The proximal base is triangular in outline. In lateral view, the base is slightly rounded in profile (distally directed concavity) and the shaft is longitudinally curved (Fig. 2c). Relative to the bone's overall length the dorsoplantar basal height is compressed, falling below the average for *Pan*, *Gorilla*, anatomically modern humans (Fig. 3e), and the single reported *Ar. ramidus* sample (see supplementary figure 4 in ref. 11). The dorsum of the BRT-VP-2/73 base does not exhibit the two 'chondral invaginations' described for *Ar. ramidus*¹¹.

Torsion along the shaft results in the long axis of the articular head being oriented about 23° medially from the dorsoplantar axis of the base. This torsion towards the hallux is on average less than that seen in *Pan* and *Gorilla*, but significantly more than that seen in anatomically

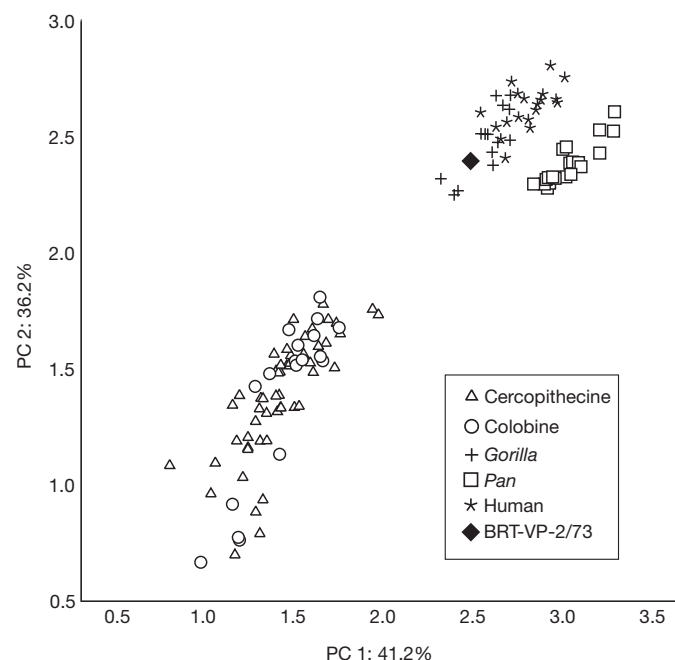


Figure 4 | Principal component analysis (PCA) of metatarsal ratios. Both PC1 and PC2 for 11 metatarsal ratios (descriptions of the ratios are provided in Supplementary Table 1) discriminate anatomically modern humans and apes from monkeys on the one hand and chimpanzees from anatomically modern humans and gorillas on the other. BRT-VP-2/73 falls in the human/gorilla cluster. Both components are heavily influenced by ratios 6, 9 and 10, which are all associated exclusively with dimensions of the hallux (see Supplementary Information for further discussion).

modern humans (on average this torsion ranges from neutral to 1–5 degrees (Supplementary Fig. 2)). In dorsal view (Fig. 2c), the shaft curves towards the medial side of the foot, a feature that in combination with the aforementioned axial torsion acts to further direct the articular head towards the hallux. This complex is characteristic of extant African apes and *Ar. ramidus* and is indicative of a grasping great toe.

In contrast to the hallucal metatarsal, the superior surface morphology of the articular head of the second metatarsal does conform to the 'typical' morphological pattern shared by *Ardipithecus* and *Australopithecus*^{6,12,16}. In distal view, the head is roughly triangular in shape and its rounded dorsal apex is domed above the epiphyseal junction. This dorsal doming creates the distinctive transverse gutter between the subchondral margin and the diaphysis, indicating the passive hyperdorsiflexion at the metatarsophalangeal joint that occurs during bipedal heel-off through toe-off^{6,11,16}.

The articulating second proximal (BRT-VP-2/73e) and intermediate (BRT-VP-2/73h) phalanges also exhibit the derived anatomical features shared with *Ar. ramidus*¹¹ and *Au. afarensis*^{6,12}. The base of the proximal phalanx exhibits the dorsiflexive anterior cant (100°) conforming to the dorsiflexion dome of the associated metatarsal head. Like *Ar. ramidus* and *Au. afarensis*, the shaft exhibits strong curvature, although not as much as in chimpanzees (Fig. 2e and Supplementary Fig. 4b; see also Supplementary Information for angle measurement methods and discussions). The inclination of the proximal articular surface in combination with the bone's longitudinal curvature results in the characteristic sulcus on the dorsal surface where the articular surface joins the shaft. The second intermediate phalanx (BRT-VP-2/73h) is relatively long compared to the associated proximal phalanx (Supplementary Table 3).

The third ray (BRT-VP-2/73f) is represented only by an isolated metatarsal head (see Fig. 2a). It too conforms to the pattern seen in *Ardipithecus*¹¹ and *Australopithecus*⁶ (Supplementary Fig. 5) in exhibiting dorsal doming. The dorsoplantar height of the third metatarsal head exceeds that of the second metatarsal, a relationship more common in *Pan*, *Gorilla* and *Australopithecus* than in anatomically modern humans wherein the second metatarsal head is usually taller (Supplementary Table 2).

The fourth ray is represented by a complete metatarsal (BRT-VP-2/73g) and its associated proximal phalanx (BRT-VP-2/73d; Fig. 2e, f). A ratio of the estimated dorsoplantar height of the metatarsal base and the bone's length indicates that the fourth metatarsal does not have the expanded, stabilizing base morphology seen in *Au. afarensis*¹² and *Homo* but, rather, is similar to *Pan* and some Old World monkeys (Supplementary Fig. 6).

The most unexpected feature seen in the fourth metatarsal is its relative length when compared to the associated first and second metatarsals. The fourth metatarsal is absolutely longer than is the second metatarsal, a condition not previously encountered in extant apes or hominins. The fourth metatarsal is also much longer than is the hallucal metatarsal and in this ratio, the fossil specimen again fails to align with extant apes or hominins and is most similar to Old World monkeys (Fig. 3d, f). At present, no associated fossil elements allow a similar comparison in *Ardipithecus* or *Australopithecus* and, as a consequence, no judgment can be reliably made regarding the polarity of this character. A relatively longer fourth metatarsal is the usual condition in Old World monkeys and it also occurs in some Miocene apes (KNM-RU 2036; ref. 17), indicating that it probably represents the primitive condition.

The proximal phalanx of the fourth ray (BRT-VP-2/73d; Fig. 2e) is well preserved and similar to those observed in *Ardipithecus*¹¹ and *Australopithecus*^{6,16}. It has the shallow transverse sulcus where the proximal articular surface cants anteriorly into the curvature of the shaft. It presents a higher degree of dorsal canting than does the phalanx of the second ray (104°, see Supplementary Fig. 4a and Supplementary Information for discussions).

Implications for hominin pedal evolution

Comparisons with the earlier *Ar. ramidus* and contemporaneous *Au. afarensis* provide a morphological and chronological context within which to view BRT-VP-2/73. Several relevant pedal elements are also represented in the South African samples from Sterkfontein and Swartkrans¹⁰ (see Supplementary Information for further discussion).

The earlier *Ar. ramidus* pedal remains indicate a mosaic foot capable of terrestrial bipedal toeing-off on the lateral four metatarsophalangeal joints (oblique metatarsal axis¹¹) while still maintaining a functionally abductable, grasping hallux. By contrast, the foot of *Au. afarensis* possessed a longitudinal pedal arch^{6,12,18}, a permanently adducted great toe^{12,14,16}, dorsal doming of its hallucal head^{6,14}, anteriorly canted bases on its proximal phalanges^{6,16} (also shared by *Ar. ramidus*¹¹), and clearly used a human-like transverse metatarsal axis during the latter stages of toe-off.

Although BRT-VP-2/73 is contemporaneous with *Au. afarensis* at around 3.4 Myr ago (see Fig. 5), it differs significantly from the known feet of *Australopithecus*. Its hallux is short and the hallucal metatarsal head lacks dorsal doming. The bases of the second and fourth metatarsals of BRT-VP-2/73 do not have the expanded dorsoplantar dimensions seen in *Ardipithecus*¹¹ and *Australopithecus*^{12,14}, features that along with the associated rugose ligamentous attachments would resist midtarsal and tarsometatarsal dorsiflexion and midfoot breaking^{19–23}. However, its lateral metatarsophalangeal joints (MTs 2, 3 and 4) do conform morphologically to the *Ardipithecus* and *Australopithecus* pattern, in having dorsally domed heads and an anterior cant to the phalangeal bases.

BRT-VP-2/73 also resembles *Ar. ramidus* in combining an abducent hallux and the medially directed torsion of the second metatarsal. It is also similar to *Ar. ramidus* and *Au. afarensis* in the metatarsophalangeal joints of the other rays, indicating that these adaptations in the lateral foot are among the earliest anatomical modifications to hominin terrestrial bipedality. The height of the hallucal metatarsal base suggests that a well-developed transverse pedal arch preceded the development of a permanent longitudinal arch. However, the lack of dorsoplantar expansion of the metatarsal bases (MTs 2 and 4) suggests that this midtarsal stabilizing feature seen in both *Ar. ramidus*¹¹ and *Au. afarensis*¹² was absent in this specimen.

The most surprising feature observed in the BRT-VP-2/73 forefoot is the length of the fourth metatarsal relative to the first and second metatarsals. The currently available *Ardipithecus* and *Australopithecus* (eastern and South African) fossil record is not adequate to assess accurately the significance of this particular feature. However, in light of its occurrence in some Miocene apes (for example, KNM-RU 2036) it may represent the primitive state in early hominins. Nonetheless, it is clear that the BRT-VP-2/73 foot skeleton represents a hominin that, unlike the contemporaneous *Au. afarensis*, retained a grasping capacity that would allow it to exploit arboreal settings more effectively. Yet, judging from its lateral metatarsophalangeal complex, when on the ground it was at least facultatively bipedal, although it may have practiced bipedality in a novel fashion probably similar to *Ar. ramidus*. Unlike *Au. afarensis*, it did not have a longitudinal pedal arch, nor was it capable of efficiently using the transverse metatarsal axis.

Although the taxonomic affinity of BRT-VP-2/73 is currently indeterminate, there is adequate morphological evidence that it does not belong to the contemporaneous species *Au. afarensis*. Regardless of its taxonomic affinity, however, this specimen is the first strong evidence indicating multiple hominin lineages, adaptively separated (at least in the foot skeleton), in the 3–4-Myr-ago time interval. A final, but important, note for the metatarsal ratios used in the PCA performed in this study, anatomically modern humans and gorillas overlap substantially and BRT-VP-2/73 falls in the gorilla cluster. It is unclear at this point what the functional implications of this overlap might mean; it requires further investigation as it has important consequences for the interpretation of locomotor behaviour in early hominins.

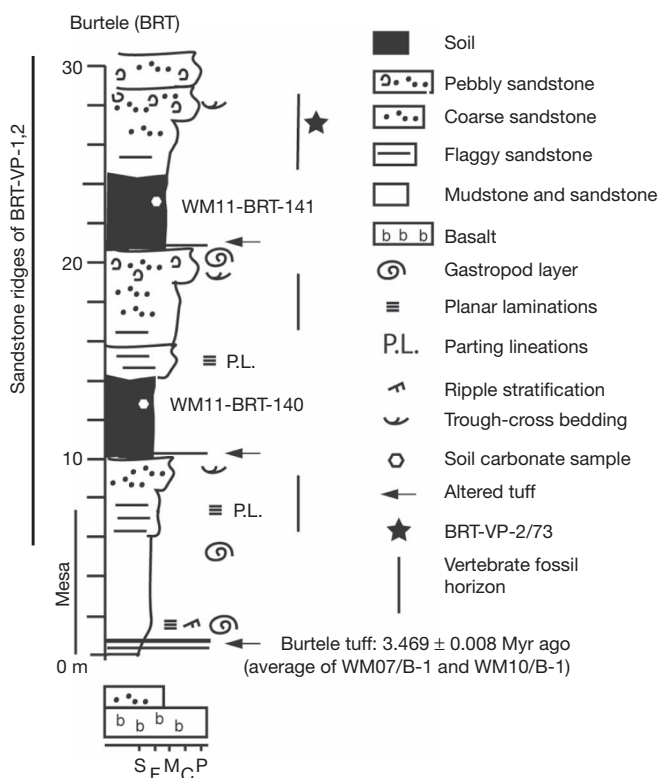


Figure 5 | Stratigraphic section at the BRT localities and placement of the BRT-VP-2/73 partial foot skeleton. The Burtele tuff is dated by the $^{40}\text{Ar}/^{39}\text{Ar}$ method to 3.469 ± 0.008 Myr ago and lies a maximum of about 27 m below BRT-VP-2/73, providing a maximum age constraint of ~ 3.47 Myr ago for the foot specimen (shown by the black star) and for three fossiliferous sandstone horizons (shown by vertical lines) at BRT-VP-1 and BRT-VP-2. An approximate age for the foot specimen, using regional sediment accumulation rates, suggests an age of between 3.2 and 3.4 Myr ago for BRT-VP-2/73 (see Methods for details). S, F, M, C, P indicates soil, flaggy, mudstone, coarse and pebbly sandstone, respectively; it shows the degree of resistance to erosion and rock stiffness.

METHODS SUMMARY

The Burtele tuff at the base of the section is dated by the $^{40}\text{Ar}/^{39}\text{Ar}$ method to 3.469 ± 0.008 Myr ago (analytical data are given in Supplementary Information) and lies a maximum of about 27 m below BRT-VP-2/73, providing a firm maximum age constraint of ~ 3.47 Myr ago for the foot specimen (Fig. 5). An approximate age for the foot specimen can be estimated using regional sediment accumulation rates. The average rate for older WORMIL strata in the Waki-Mille confluence area is 11 cm kyr^{-1} (ref. 2), which yields an estimated age of 3.22 Myr ago for the BRT-VP-2/73 specimen. This rate is much lower than estimates for the Sidi Hakoma Member of the Hadar Formation^{24–26}, which is closer in age to the BRT ridge section, but is much farther away geographically. Using a Sidi Hakoma accumulation rate of 30 cm kyr^{-1} yields an estimate of 3.38 Myr ago for BRT-VP-2/73. These contrasting rates indicate an age of between 3.2 and 3.4 Myr ago for BRT-VP-2/73.

For the isotopic analysis of pedogenic carbonate, carbonate nodules were sampled from peds with slickenside surfaces and clay cutans, within a distinct pedogenic carbonate zone, $\geq 50 \text{ cm}$ below the palaeosol contact with the overlying silt. $\delta^{13}\text{C}$, $\delta^{18}\text{O}$ and Δ_{47} measurements of carbonate were made using an automated common acid bath peripheral coupled to a Thermo MAT 253 mass spectrometer at Johns Hopkins University, using methods described previously²⁷. The results are reported in Supplementary Table 8.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 October 2011; accepted 8 February 2012.

- Haile-Selassie, Y., Deino, A., Saylor, B., Umer, M. & Latimer, B. Preliminary geology and paleontology of new hominid-bearing Pliocene localities in the central Afar region of Ethiopia. *Anthropol. Sci.* **115**, 215–222 (2007).

- Deino, A. L. et al. $^{40}\text{Ar}/^{39}\text{Ar}$ dating, paleomagnetism, and tephrochemistry of Pliocene strata of the hominid-bearing Woranso-Mille area, west-central Afar Rift, Ethiopia. *J. Hum. Evol.* **58**, 111–126 (2010).
- Haile-Selassie, Y., Saylor, B. Z., Deino, A., Alene, M. & Latimer, B. New hominid fossils from Woranso-Mille (Central Afar, Ethiopia) the taxonomy of early *Australopithecus*. *Am. J. Phys. Anthropol.* **141**, 406–417 (2009).
- Haile-Selassie, Y. et al. An early *Australopithecus afarensis* postcranium from Woranso-Mille, Ethiopia. *Proc. Natl Acad. Sci. USA* **107**, 12121–12126 (2010).
- Haile-Selassie, Y. Phylogeny of early *Australopithecus*: new fossil evidence from the Woranso-Mille (Central Afar, Ethiopia). *Phil. Trans. R. Soc. B* **365**, 3323–3331 (2010).
- Latimer, B. M. & Lovejoy, C. O. Hominid tarsal, metatarsal, and phalangeal bones recovered from the Hadar formation: 1974–1977 collections. *Am. J. Phys. Anthropol.* **57**, 701–719 (1982).
- Day, M. H. & Napier, J. R. Hominid fossils from Bed I, Olduvai Gorge, Tanganyika: fossil footbones. *Nature* **201**, 969–970 (1964).
- Zipfel, B. et al. The foot and ankle of *Australopithecus sediba*. *Science* **333**, 1417–1420 (2011).
- Clarke, R. J. & Tobias, P. V. Sterkfontein Member 2 foot bones of the oldest South African hominid. *Science* **269**, 521–524 (1995).
- Deloison, Y. Anatomie des os fossiles de pieds des hominides d'Afrique du Sud. *Biomet. Hum. Anthropol.* **21**, 189–230 (2003).
- Lovejoy, O. C., Latimer, B., Suwa, G., Asfaw, B. & White, T. D. Combining prehension and propulsion: The foot of *Ardipithecus ramidus*. *Science* **326**, 72 (2009).
- Ward, C. V., Kimbel, W. H. & Johanson, D. C. Complete fourth metatarsal and arches in the foot of *Australopithecus afarensis*. *Science* **331**, 750–753 (2011).
- Harcourt-Smith, W. E. H. & Aiello, L. C. Fossils, feet and the evolution of human bipedal locomotion. *J. Anat.* **204**, 403–416 (2004).
- Latimer, B. M. & Lovejoy, C. O. Hallux tarsometatarsal joint in *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* **82**, 125–133 (1990).
- Griffin, N. L. & Richmond, B. G. Joint orientation and function in great ape and human proximal pedal phalanges. *Am. J. Phys. Anthropol.* **141**, 116–123 (2010).
- Latimer, B. M. & Lovejoy, C. O. Metatarsophalangeal joints of *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* **83**, 13–23 (1990).
- Walker, A. C. & Pickford, M. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 325–413 (Plenum, 1983).
- White, T. D. & Suwa, G. Hominid footprints at Laetoli: facts and interpretations. *Am. J. Phys. Anthropol.* **72**, 485–514 (1987).
- Elftman, H. & Manter, J. Chimpanzee and human feet in bipedal walking. *Am. J. Phys. Anthropol.* **20**, 69–79 (1935).
- Elftman, H. & Manter, J. The evolution of the human foot with special reference to the joints. *J. Anat.* **70**, 56–67 (1935).
- DeSilva, J. M. Functional morphology of the ankle and the likelihood of climbing in early hominins. *Proc. Natl Acad. Sci. USA* **106**, 6567–6572 (2009).
- D'Aout, K., Aerts, P., De Clercq, D., De Meester, K. & Van Elsacker, L. Segment and joint angles of hind limb during bipedal and quadrupedal walking of the bonobo (*Pan paniscus*). *Am. J. Phys. Anthropol.* **119**, 37–51 (2002).
- Vereecke, E., D'Aout, K., De Clercq, D., Van Elsacker, L. & Aerts, P. Dynamic plantar pressure distribution during terrestrial locomotion of bonobos (*Pan paniscus*). *Am. J. Phys. Anthropol.* **120**, 373–383 (2003).
- Walter, R. C. & Aronson, J. L. Age and source of the Sidi Hakoma Tuff, Hadar Formation, Ethiopia. *J. Hum. Evol.* **25**, 229–240 (1993).
- deMenocal, P. B. & Brown, F. H. in *Hominid Evolution and Climatic Change in Europe* (eds Agustí, J., Rook, L. & Andrews, P.) 23–54 (Cambridge Univ. Press, 1999).
- McDougall, I. & Brown, F. H. Geochronology of the pre-KBS Tuff sequence, Omo Group, Turkana Basin. *J. Geol. Soc. Lond.* **165**, 549–562 (2008).
- Passy, B. H., Levin, N. E., Cerling, T. E., Brown, F. H. & Eiler, J. M. High-temperature environments of human evolution in East Africa based on bond ordering in paleosol carbonates. *Proc. Natl Acad. Sci. USA* **107**, 11245–11249 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Authority for Research and Conservation of Cultural Heritage and the Afar Regional State of Ethiopia for permission to conduct field and laboratory research, and the Afar people of the Woranso-Mille area for support in the field. We also thank M. Asnake, R. Bernor, S. Frost, D. Geraads, I. Giaourtsakis, M. Lewis, W. Sanders and L. Werdelin for faunal identifications. We thank B. Passy for aid with isotope analyses; E. Guthrie for unpublished primary data; L. Russell for photography; S. Melillo and H. Gebreyesus for fieldwork; O. Lovejoy, S. Simpson, G. Suwa and T. White for comments and discussions; D. Su for discussions and assistance in statistical analysis; and L. Jellema for assistance in photography. This research was supported by funding from the LSB Leakey Foundation, the National Geographic Society, the Cleveland Museum of Natural History, and NSF grants BCS-0234320, BCS-0321893, BCS-0542037 and BCS-1124705.

Author Contributions Y.H.-S. and B.M.L. conducted the description and comparative analysis. B.Z.S., N.E.L. and M.A. compiled the stratigraphic sequence. A.D. conducted the radiometric dating. N.E.L. conducted stable isotope analysis. Y.H.-S. and B.M.L. wrote the paper with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.H.-S. (yhailese@cmnh.org).

METHODS

$^{40}\text{Ar}/^{39}\text{Ar}$ dating procedures follow those of ref. 2. The mineral separates were irradiated for 5 h in two separate batches in the in-core CLICIT facility of the Oregon State University TRIGA reactor. Sanidine from the Fish Canyon Tuff of Colorado was used as a monitor mineral, with an age of 28.201 Myr²⁸. After irradiation, the feldspar grains were individually analysed under ultra-high vacuum on a MAP 215 Noble-gas mass spectrometer, using a focused CO_2 laser as the heating device. In all, 56 grains were analysed from the two samples (Supplementary Table 7). Most grains (34) proved to be K-feldspar, as judged by the Ca/K ratio determined from the measured argon isotopes, whereas the remainder were relatively low-Ca/K plagioclase. Most analyses yielded the anticipated high proportion of radiogenic ^{40}Ar relative to atmospheric ^{40}Ar contamination expected for unaltered feldspars from Pliocene volcanic rocks, but a few exhibited anomalously low radiogenic content, and were excluded from further analysis; an arbitrary cutoff of 60% $^{40}\text{Ar}^*$ was used, identifying four grains for exclusion. In addition, as is typical for East African tephra, a slight tail of the age distribution towards older ages was observed. A statistical filter was applied to the sample distributions, using a median outlier determinant (outliers were classified as falling 1.5 'normalized median absolute deviations' from the median). Use of this criterion identified three outliers in each of WM07/B-1 (K-feldspar) and WM10/B-1 (Plagioclase). The remaining populations yield simple, unimodal Gaussian-like distributions (Supplementary Fig. 7). Weighted-mean sample ages of the K-feldspar populations from samples WM07/B-1 and WM10/B-1 are 3.484 ± 0.011 Myr ($n = 24$; 1σ analytical error, incorporating error in J , the neutron fluence parameter of 0.2%) and 3.453 ± 0.011 Myr ($n = 4$), respectively (Supplementary Table 4). An overall weighted-mean of the two K-feldspar ages is 3.469 ± 0.008 Myr, taken as the reference age for the

Burtele tuff. The plagioclase weighted-mean age of sample WM10/B-1 is predictably less precise than either K-feldspar age, due to the lower potassium content, but is nevertheless a reasonable result (3.42 ± 0.03 Myr) that is not statistically different from the K-feldspar age.

The Burtele tuff at the base of the section is dated by the $^{40}\text{Ar}/^{39}\text{Ar}$ method to 3.469 ± 0.008 Myr ago (analytical data are given in Supplementary Information) and lies a maximum of about 27 m below BRT-VP-2/73, providing a firm maximum age constraint of ~ 3.47 Myr ago for the foot specimen. An approximate age for the foot specimen can be estimated using regional sediment accumulation rates. The average rate for older WORMIL strata in the Waki-Mille confluence area is 11 cm kyr^{-1} (ref. 2), which yields an estimated age of 3.22 Myr ago for the BRT-VP-2/73 specimen. This rate is much lower than estimates for the Sidi Hakoma Member of the Hadar Formation^{24–26}, which is closer in age to the BRT ridge section, but is much farther away geographically. Using a Sidi Hakoma accumulation rate of 30 cm kyr^{-1} yields an estimate of 3.38 Myr for BRT-VP-2/73. These contrasting rates suggest an age of between 3.2 and 3.4 Myr ago for BRT-VP-2/73.

For the isotopic analysis of pedogenic carbonate, carbonate nodules were sampled from peds with slickenside surfaces and clay cutans, within a distinct pedogenic carbonate zone, ≥ 50 cm below the palaeosol contact with the overlying silt. $\delta^{13}\text{C}$, $\delta^{18}\text{O}$ and Δ_{47} measurements of carbonate were made using an automated common acid bath peripheral coupled to a Thermo MAT 253 mass spectrometer at Johns Hopkins University, using methods described previously²⁷. The results are reported in Supplementary Table 8.

28. Kuiper, K. F. *et al.* Synchronizing rock clocks of earth history. *Science* **320**, 500–504 (2008).

Systematic identification of genomic markers of drug sensitivity in cancer cells

Mathew J. Garnett^{1*}, Elena J. Edelman^{2*}, Sonja J. Heidorn^{1*}, Chris D. Greenman^{1†}, Anahita Dastur², King Wai Lau¹, Patricia Greninger², I. Richard Thompson¹, Xi Luo², Jorge Soares¹, Qingsong Liu^{3,4}, Francesco Iorio^{1,5}, Didier Surdez⁶, Li Chen², Randy J. Milano², Graham R. Bignell¹, Ah T. Tam², Helen Davies¹, Jesse A. Stevenson², Syd Barthorpe¹, Stephen R. Lutz², Fiona Kogera¹, Karl Lawrence¹, Anne McLaren-Douglas¹, Xenia Mitropoulos², Tatiana Mironenko¹, Helen Thi², Laura Richardson¹, Wenjun Zhou^{3,4}, Frances Jewitt¹, Tinghu Zhang^{3,4}, Patrick O'Brien¹, Jessica L. Boisvert², Stacey Price¹, Wooyoung Hur^{3,4}, Wanjuan Yang¹, Xianming Deng^{3,4}, Adam Butler¹, Hwan Geun Choi^{3,4}, Jae Won Chang^{3,4}, Jose Baselga², Ivan Stamenkovic⁷, Jeffrey A. Engelman², Sreenath V. Sharma^{2†}, Olivier Delattre⁶, Julio Saez-Rodriguez⁵, Nathanael S. Gray^{3,4}, Jeffrey Settleman², P. Andrew Futreal¹, Daniel A. Haber^{2,8}, Michael R. Stratton¹, Sridhar Ramaswamy², Ultan McDermott¹ & Cyril H. Benes²

Clinical responses to anticancer therapies are often restricted to a subset of patients. In some cases, mutated cancer genes are potent biomarkers for responses to targeted agents. Here, to uncover new biomarkers of sensitivity and resistance to cancer therapeutics, we screened a panel of several hundred cancer cell lines—which represent much of the tissue-type and genetic diversity of human cancers—with 130 drugs under clinical and preclinical investigation. In aggregate, we found that mutated cancer genes were associated with cellular response to most currently available cancer drugs. Classic oncogene addiction paradigms were modified by additional tissue-specific or expression biomarkers, and some frequently mutated genes were associated with sensitivity to a broad range of therapeutic agents. Unexpected relationships were revealed, including the marked sensitivity of Ewing's sarcoma cells harbouring the *EWS* (also known as *EWSR1*)-*FLI1* gene translocation to poly(ADP-ribose) polymerase (PARP) inhibitors. By linking drug activity to the functional complexity of cancer genomes, systematic pharmacogenomic profiling in cancer cell lines provides a powerful biomarker discovery platform to guide rational cancer therapeutic strategies.

There is compelling evidence that the likelihood of a patient's cancer responding to treatment can be strongly influenced by alterations in the cancer genome. For example, the use of drugs to selectively target the protein product of the *BCR-ABL* translocation in chronic myeloid leukaemia (CML) has revolutionized the treatment of this disease, with five-year survival rates of 90% in treated patients¹. Although targeting of specific genetic changes in defined patient subsets has been successful, a poorly explained range of responses to appropriately selected therapies is often still observed in patients^{2,3}. Moreover, a large number of cancer drugs have not been linked to specific genomic alterations that could be used as biomarkers to specify their selective therapeutic effectiveness⁴. As drug pipelines generate new classes of compounds, systematic methods to identify predictive biomarkers during their early development could have a profound effect on the design, cost and ultimate success of new cancer drug development.

The NCI60 cell line panel and associated drug screens pioneered the approach of using cancer cell lines to link drug sensitivity with genotype data^{5,6}. Cancer cell lines have subsequently been used to identify rare drug-sensitizing genotypes, including mutant *EGFR*, *BRAF* and the *EML4-ALK* translocation, which are highly predictive of clinical responses^{2,3,7}. Here, we report the results of a large-scale screen of human cancer cell lines, incorporating detailed genomic and gene expression analysis, to identify systematically drug-sensitivity biomarkers to a broad range of cancer drugs.

Therapeutic biomarker discovery

To capture the high degree of genomic diversity in cancer and to identify rare mutant subsets with altered drug sensitivity, we assembled 639 human tumour cell lines, representing the spectrum of common and rare types of adult and childhood cancers of epithelial, mesenchymal and haematopoietic origin (Fig. 1a and Supplementary Data 1). Cell lines were subjected to sequencing of the full coding exons of 64 commonly mutated cancer genes, genome-wide analysis of copy number gain and loss using Affymetrix SNP6.0 microarrays, and expression profiling of 14,500 genes using Affymetrix HT-U133A microarrays. The presence of seven commonly rearranged cancer genes and of microsatellite instability (MSI) was also investigated. The 130 drugs selected for screening covered a wide range of targets and processes implicated in cancer biology (Fig. 1b and Supplementary Data 2). They encompassed both targeted agents ($n = 114$) and cytotoxic chemotherapeutics ($n = 13$), including approved drugs used in clinical practice ($n = 31$), drugs in development undergoing studies in clinical trials ($n = 47$), and experimental tool compounds ($n = 52$). To gain insight into drug-to-drug variation, we included multiple drugs designed against well-credentialed targets (Fig. 1b). The effect of 72 h of drug treatment on cell viability was used to derive a multi-parameter description of drug sensitivity, including the half-maximal inhibitory concentration (IC_{50}), and the slope of the dose-response curve (Supplementary Fig. 1). In total, we assayed 48,178 drug-cell-line

¹Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. ²Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, Massachusetts 02129, USA.

³Department of Cancer Biology, Dana Farber Cancer Institute, 44 Binney Street, Boston Massachusetts 02115, USA. ⁴Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, Massachusetts 02115, USA. ⁵EMBL-EBI, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK. ⁶Laboratoire de génétique et biologie des cancers, Institut Curie, 75248 Paris, Cedex 05, France. ⁷Division of Experimental Pathology, Institute of Pathology, Centre Hospitalier Universitaire Vaudois (CHUV), 1005 Lausanne, Switzerland. ⁸Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. [†]Present addresses: Department of Computing, University of East Anglia, Norwich NR4 7TJ, UK (C.D.G.); The Genome Analysis Centre, Norwich Research Park, Norwich NR4 7UH, UK (C.D.G.); Oncology Drug Discovery, Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA (S.V.S.).

*These authors contributed equally to this work.

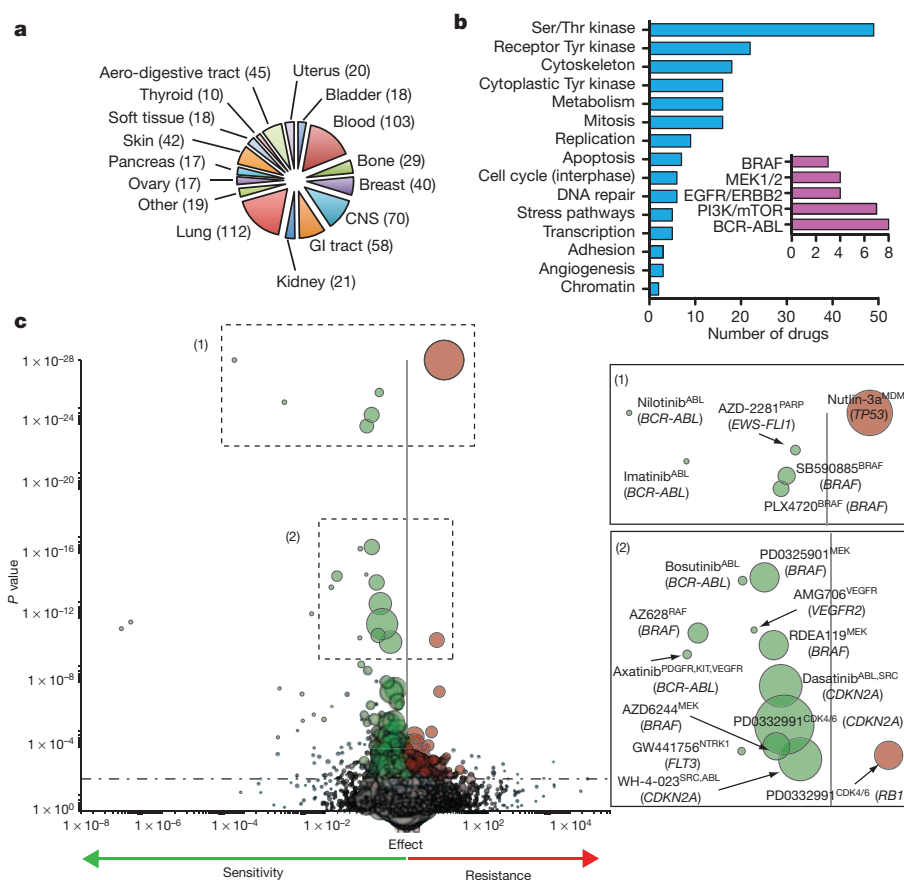


Figure 1 | A systematic screen in cancer cell lines identifies therapeutic biomarkers. **a**, The number of tumour-derived cell lines used for screening classified according to tissue type ($n = 639$ in total). CNS, central nervous system; GI, gastrointestinal. **b**, The panel of 130 screening drugs classified according to their therapeutic targets, primary effector pathways, and cellular functions. A single drug may be included in multiple categories. The inset indicates the number of drugs screened against a selection of prototype cancer targets. **c**, A volcano plot representation of MANOVA results showing the magnitude (effect, x -axis) and significance (P value, inverted y -axis) of all

drug-gene associations. Each circle represents a single drug-gene interaction and the size is proportional to the number of mutant cell lines screened (range 1–334). The horizontal dashed line indicates the threshold of statistical significance (0.2 false discovery rate, $P < 0.0099$). Insets (1) and (2) are magnified views of selected highly significant associations; the drug name, therapeutically relevant target(s) (in superscript) and cancer gene (in brackets) are given for each. The P values for nilotinib^{ABL} (BCR-ABL), $P = 2.54 \times 10^{-65}$, and nutlin-3a^{MDM2} (TP53), $P = 2.78 \times 10^{-37}$, have been capped at 1×10^{-28} in this representation.

combinations with a range of 275–507 cell lines screened per drug (mean = 368 cell lines per drug; Supplementary Data 2 and Supplementary Fig. 2). Clustering of compounds across cell lines based on IC_{50} values indicated that drugs with overlapping specificity were highly correlated, supporting the selectivity of the biological effects observed in the data set (Supplementary Fig. 3 and Supplementary Tables 1 and 2).

Tumours from a particular tissue frequently have a shared set of somatic mutations. To gain insight into how this might relate to drug sensitivity, we performed an analysis to identify associations between cancer tissue type and drug sensitivity based on IC_{50} values. As expected, in some instances tumour-type-specific sensitivity may be explained by the prevalence of cancer gene mutations (for example, breast cancer sensitivity to inhibitors of the phosphatidylinositol 3-kinase (PI3K) pathway that is commonly altered in this tumour type; Supplementary Data 3 and 4). In other cases, however, our current understanding of cancer genomes could not explain the observed associations. For instance, renal cell carcinoma (RCC) cells were sensitive to five SRC inhibitors (for example, AZD0530, $P < 1 \times 10^{-4}$, $n = 9$ RCC and 294 non-RCC cell lines)⁸, glioma cells were sensitive to a ROCK inhibitor (GSK269962A, $P < 1 \times 10^{-6}$, $n = 23$ glioma and 266 non-glioma cell lines)⁹. This analysis also identified therapeutic associations already used in the clinic with incompletely understood molecular bases, such as the sensitivity of myeloma cells to lenalidomide ($P < 1 \times 10^{-5}$, $n = 3$ myeloma and

455 non-myeloma cell lines)¹⁰. For most drugs, however, sensitive cell lines were scattered across multiple cancer types.

Cancer genes are drug-sensitivity biomarkers

Single gene mutations are increasingly being adopted as clinical biomarkers for the optimal application of cancer therapeutics. To identify associations between individual mutated cancer genes and drug sensitivity across the cell line panel, we used a multivariate analysis of variance (MANOVA) incorporating the IC_{50} value and slope of the dose-response curve. This analysis revealed a large number of individual gene-drug associations, a subset of which (448/9,039, 5%) were highly significant and are discussed here (Fig. 1c and Supplementary Data 5). Interestingly, most of the cancer genes analysed, including those that are not known to be direct targets of the drugs tested, were associated with either sensitivity or resistance to at least one drug in our panel (65/69, 94%) (Supplementary Fig. 4). Similarly, sensitivity to most drugs tested was associated with a mutation in at least one cancer gene (118/130, 91%). Thus, diverse cancer gene mutations are implicated as markers of sensitivity or resistance to a broad range of anticancer drugs, indicating that genomic biomarkers could inform the therapeutic selectivity of many cancer drugs.

The mutated cancer genes most clearly associated with drug sensitivity are oncogenes that are direct targets of the relevant drug. For example, the BCR-ABL rearrangement conferred sensitivity to multiple ABL inhibitors (for example, $P = 2.54 \times 10^{-65}$ for nilotinib;

Fig. 1c and Supplementary Fig. 5)¹, several of which are approved for CML treatment. Similarly, *BRAF* mutation was associated with sensitivity to BRAF and inhibitors of MEK1 and MEK2 (for example, $P = 1.25 \times 10^{-24}$ for PLX4720; Figs 1c and 2a–c)³, including a structural analogue of vemurafenib, which in clinical trials has extended the survival of *BRAF*-mutation-positive melanoma patients. Additionally, *ERBB2* (also known as *HER2*) amplification was associated with sensitivity to EGFR-family inhibitors including lapatinib ($P < 1 \times 10^{-7}$; Fig. 2d)¹¹, which is licensed for the treatment of HER2-positive breast cancer. We were also able to detect known associations between *EGFR*, *FLT3* and *PIK3CA* mutations and drugs that target the products of these genes (Supplementary Data 5)^{12,13}. A number of associations were driven by marked responses in small subsets of outlier cell lines. For example, two *FGFR2*-mutated cell lines were exquisitely sensitive to the FGFR inhibitor PD-173074 (Fig. 2e; $P < 1 \times 10^{-5}$)^{14,15}, confirming the need for large panels of cell lines to capture low-frequency drug-sensitizing genotypes.

We also found associations between the presence of inactivating mutations in tumour suppressor genes and several drugs, which in some instances provide insight into the interplay between tumour suppressors and the cellular machinery in mediating drug sensitivity.

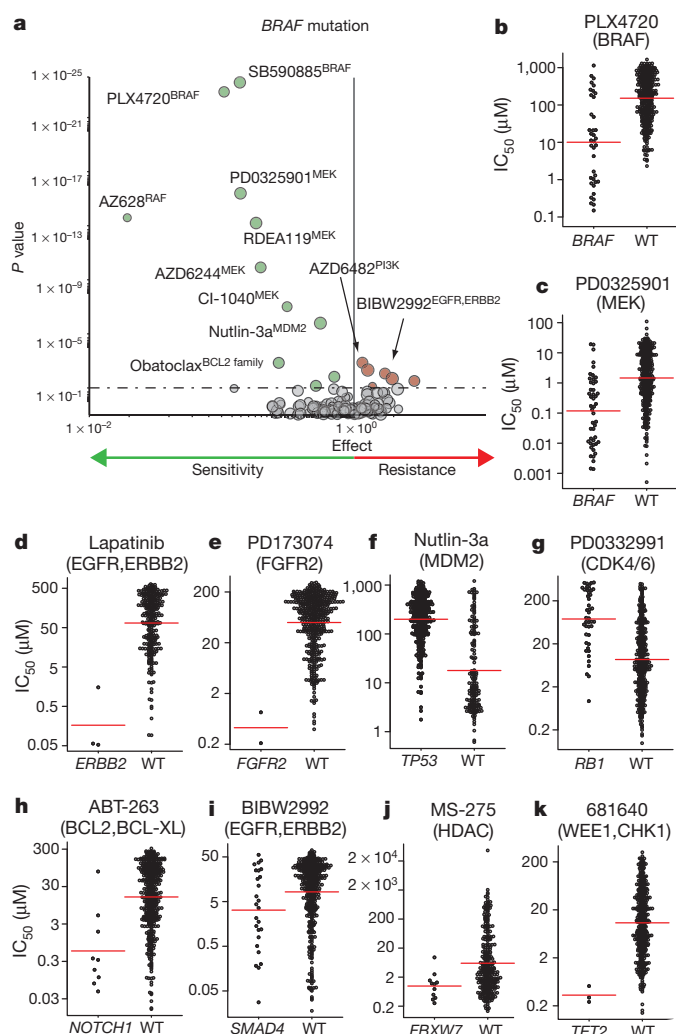


Figure 2 | Biomarkers of drug sensitivity and resistance. a, Gene-specific volcano plot of drug sensitivity associated with *BRAF* mutations in cancer cell lines (range 22–54). b–k, Scatter plots of cell line IC_{50} (μM) values from selected drug–gene associations. IC_{50} values are on a log scale comparing mutated (gene symbol given) or non-mutated (wild type (WT)) cell lines. Each circle represents the IC_{50} of one cell line and the red bar is the geometric mean. The drug name is indicated above each plot and therapeutic drug target(s) are bracketed.

For example, mutation of *TP53*, an important regulator of apoptosis and cell cycle arrest in response to cellular stress, confers resistance to nutlin-3a ($P < 1 \times 10^{-36}$), an inhibitor of the MDM2 E3 ligase that negatively regulates p53 protein levels (Fig. 2f)¹⁶. Similarly, mutational inactivation of *RBI1*, a key repressor of cell cycle progression in normal cells, confers resistance to PD-0332991 ($P < 1 \times 10^{-10}$), an inhibitor of the upstream cyclin-dependent kinases (CDKs) 4 and 6, which drive cell cycle progression by inhibiting pRb through phosphorylation (Fig. 2g)¹⁷. Conversely, mutational inactivation of *CDKN2A*, encoding the CDK inhibitory protein p16, was associated with sensitivity to PD-0332991 ($P < 1 \times 10^{-11}$; Supplementary Data 5)¹⁷, presumably because *CDKN2A*-mutated cells have an enhanced requirement for signalling through the CDK4/6–pRb signalling pathway.

In other instances genomic associations appear related to enrichment of mutations in a particular tissue type. The association of the *BRAF* and *NRAS* mutations with sensitivity to obatoclax mesylate, a pro-apoptotic drug that targets BCL2 family anti-apoptotic proteins (BCL2, BCL-XL (also known as BCL2L1) and MCL1), probably results from the enrichment of these mutations in melanoma, as drug sensitivity among melanoma cell lines was not correlated with the presence or absence of these mutations (Supplementary Fig. 6). The tissue-specific effect of obatoclax may be related to inhibition of the melanoma survival mediator MCL1 (ref. 18), because sensitivity of melanoma lines to ABT-263, another BCL2 inhibitor that does not target MCL1, was not correlated with *BRAF* or *NRAS* mutation. Moreover, an ABT-263-insensitive melanoma cell line can be sensitized to this drug by short interfering (si)RNA-mediated depletion of MCL1 (Supplementary Fig. 7).

The genomic associations identified for 13 clinically approved cytotoxic chemotherapeutics in our panel were generally less significant than for targeted drugs, indicating that single gene biomarkers may be less informative for this class of drugs with broad action across many cancers (Supplementary Figs 8 and 9). Intriguingly, we did not find general associations between targeted or cytotoxic drug-sensitivity patterns and mutations in *TP53*. It may be that functional inactivation of p53, through mutations or abrogation of signalling pathways that regulate its activity, is an almost universal feature of cancer cell lines and thus differential drug sensitivity between mutant and non-mutant cell lines is not observed¹⁹.

Several other novel gene–drug associations were identified that cannot be readily explained on the basis of our current knowledge of signalling pathways and may reflect unappreciated biological relationships. Mutation of *NOTCH1* was associated with sensitivity to ABT-263 ($P < 1 \times 10^{-9}$; Fig. 2h and Supplementary Fig. 10), perhaps due to decreased expression of BCL2 family members in *NOTCH1*-mutant cell lines (Supplementary Fig. 11). Amplification of *CCND1* (cyclin D1) or loss of *SMAD4* were associated with sensitivity to multiple EGFR-family inhibitors including lapatinib and BIBW2992; and for *SMAD4* this correlated with elevated *EGFR* gene expression (Fig. 2i and Supplementary Fig. 12). Inactivation of *STK11* (also known as *LKB1*; $P < 0.01$), thought to relieve repression of mTOR, was associated with sensitivity to the HSP90 inhibitor 17-AAG. Additionally, loss of *FBXW7* was associated with sensitivity to the histone deacetylase (HDAC) inhibitor MS-275 ($P < 1 \times 10^{-5}$; Fig. 2j), and *TET2* loss was associated with sensitivity to the WEE1 and CHK1 (also known as CHEK1) inhibitor 681640 ($P < 1 \times 10^{-4}$; Fig. 2k). These associations, and others presented here (Supplementary Data 5), represent candidate biomarkers of drug sensitivity and may ultimately be useful for the deployment of targeted therapies in cancer.

Complex genomic correlates of drug sensitivity

In most instances sensitivity of cancer cells to drugs is likely to depend on a multiplicity of genomic and epigenomic variables. Indeed, single gene–drug associations were only rarely able to explain the range of drug sensitivities observed across cell lines for any given drug (Fig. 2). We thus applied elastic net regression²⁰, a penalized linear modelling

technique, to identify cooperative interactions among multiple genes and transcripts across the genome and defined response signatures for each drug. Elastic net regression identified 26,938 feature–drug associations (Supplementary Data 6) from which 534 associations corresponding to 69 different drugs were highly significant (defined as $-2.95 > \text{effect } (e) > 2.79$ and frequency $(f) > 0.76$; Fig. 3a and Supplementary Fig. 13 and Supplementary Data 7).

In many instances transcriptional features showed correlations with drug sensitivity that were equal to or stronger than those observed with gene mutation (Fig. 3a and Supplementary Table 3). For example, although sensitivity to the EGFR and ERBB2 inhibitor lapatinib correlated with *ERBB2* expression and mutation, the strongest correlate for this drug was actually expression of the matrix metalloproteinase *MMP28* ($e = -29.28, f = 1$) (Fig. 3a). Notably, for most drugs, including

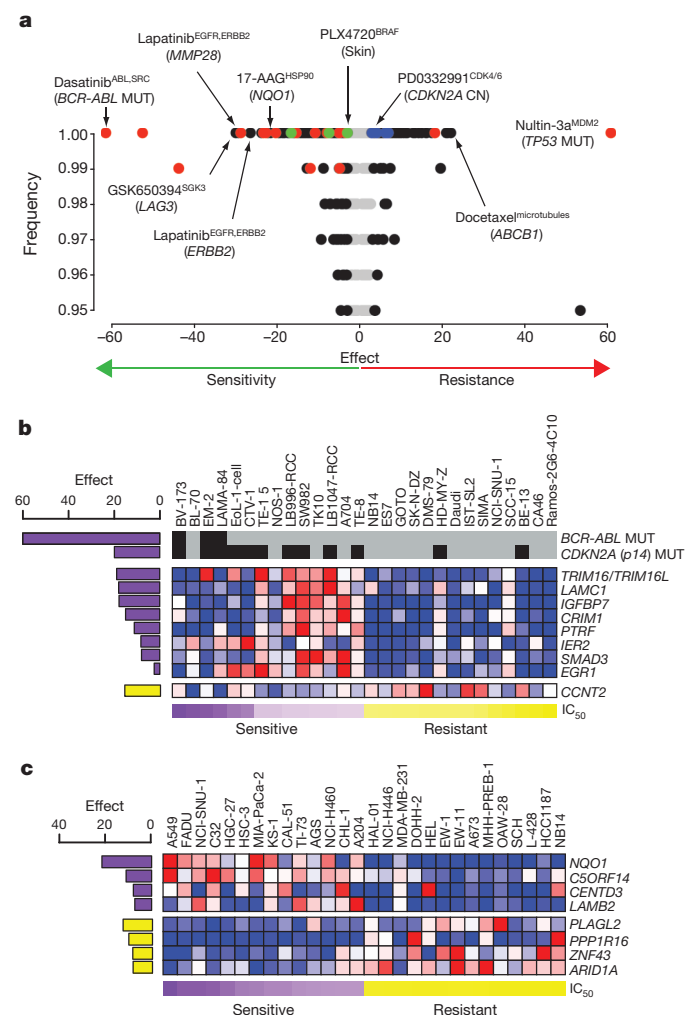


Figure 3 | Multi-feature genomic signatures of drug response. **a**, The top drug-feature associations identified by the elastic net are plotted for their frequency and effect size. Associations are coloured black for expression features, red for mutations, blue for copy number, and green for tissue. MUT, mutation. CN, copy number. **b, c**, Heatmaps of highly significant elastic net features associated with response to dasatinib (inhibitor of SRC and ABL) (**b**) and 17-AAG (HSP90 inhibitor) (**c**) for the 14 most sensitive (purple) and resistant (yellow) cell lines. For each cell line mutation features are at the top of the heatmap shown in black (present) or grey (absent), followed by expression features (blue corresponds to lower expression, red to higher expression). To the left of each feature is a bar indicating the absolute value of the effect size. Bars in purple are negative effects, indicating features associated with sensitivity, and bars in yellow are positive effects, indicating features associated with resistance. The natural log IC_{50} values are represented at the bottom. For clarity, only the top four features associated with sensitivity and resistance to 17-AAG are shown.

those with clear linkage to cancer gene mutations, elastic net modelling identified multi-feature signatures of drug sensitivity. For example, together with *BRAF* mutation, sensitivity to RAF or MEK1 and MEK2 inhibitors was recurrently associated with 67 features. These features included expression of *SPRY2*, *DUSP4* and *DUSP6*, which are known regulators of MAPK signalling (Supplementary Fig. 14)^{21,22}. Interestingly, expression of 8 genes identified as markers of sensitivity to the MEK inhibitor AZD6244 significantly overlapped with an 18-gene signature of sensitivity to this drug (hypergeometric test of the overlap significance: $P = 3 \times 10^{-9}$)²³. In some cases, elastic net modelling identified complex patterns of biomarkers corresponding to distinct subsets of sensitive cancer cell lines. Thus, sensitivity to dasatinib, an inhibitor of multiple kinases including ABL and SRC, correlated with both *BCR-ABL* translocation and with multi-gene transcriptional signatures that were expressed in sensitive cell lines lacking that gene translocation (Fig. 3b).

Elastic net modelling also identified transcriptional correlates of sensitivity to drugs without a known sensitizing mutational event. This included expression of *LAG3*, which correlated with sensitivity to the SGK inhibitor GSK-650394 ($e = -29.9, f = 1$) and the correlation between expression of the NADPH dehydrogenase family member *NQO1* with sensitivity to the HSP90 inhibitor 17-AAG ($e = -22.21, f = 1$; Fig. 3c). Consistent with these findings, *NQO1* was previously shown to metabolize 17-AAG into a more potent inhibitor of HSP90 (ref. 24).

A small number of features were recurrently associated with altered sensitivity to drugs from different classes, indicating that they may be broadly involved in mediating drug sensitivity by diverse mechanisms such as drug efflux (for example, *ABCB1*; Supplementary Data 6). To give further insight into this data set, we mapped the elastic net drug signatures onto the target of the drugs (Supplementary Data 7) and onto 457 known cancer genes (<http://www.sanger.ac.uk/genetics/CGP/Census/>; Supplementary Data 8). Collectively, these observations illustrate that in many instances multi-feature genomic signatures incorporating markers related to mutations, tissue lineage, cellular differentiation states and cellular pathways have the potential to expand and refine our current understanding of drug sensitivity.

EWS-FLI1 is a PARP inhibitor biomarker

We identified a highly significant association between the *EWS-FLI1* rearrangement that is characteristic of Ewing's sarcoma tumours and sensitivity to olaparib (AZD2281), an inhibitor of PARP ($P = 1.03 \times 10^{-26}$; Figs 1c and 4a). Screening of a structurally distinct PARP inhibitor, AG-014699, across a large panel of cell lines confirmed the sensitivity of Ewing's sarcoma cell lines (geometric mean IC_{50} for *EWS-FLI1* = 4.7 μ M versus 64 μ M for wild type, $P < 0.0001$ Mann-Whitney test, $n = 291$ cell lines; Fig. 4a and Supplementary Fig. 15 and Supplementary Data 9). Cells from Ewing's sarcoma were more sensitive to olaparib ($P = 2.84 \times 10^{-11}$) than cells from other tumour types, including sarcomas of bone and soft tissue (Supplementary Fig. 15 and Supplementary Data 3). PARP inhibitors have activity in *BRCA1*- and *BRCA2*-mutant cancers due to the defects in homologous recombination present in these tumours and their consequent reliance on alternative DNA damage repair pathways that are targeted by these inhibitors²⁵. A comparison of olaparib and AG-014699 sensitivity in a panel of cell lines using a 6-day viability assay and colony formation experiments confirmed the marked sensitivity of Ewing's sarcoma to PARP inhibitors, an effect that was comparable to that observed in *BRCA*-deficient cells (Fig. 4b, c and Supplementary Figs 16 and 17)²⁶. Furthermore, treatment with olaparib selectively induced apoptosis in Ewing's sarcoma compared to control cells (Fig. 4d). Unlike in Ewing's sarcoma, we did not observe an association between *BRCA1* and *BRCA2* mutations and sensitivity to PARP inhibitors in the 3-day screening format, which is probably due to a requirement for several rounds of division in these cells to accumulate toxic levels of DNA damage.

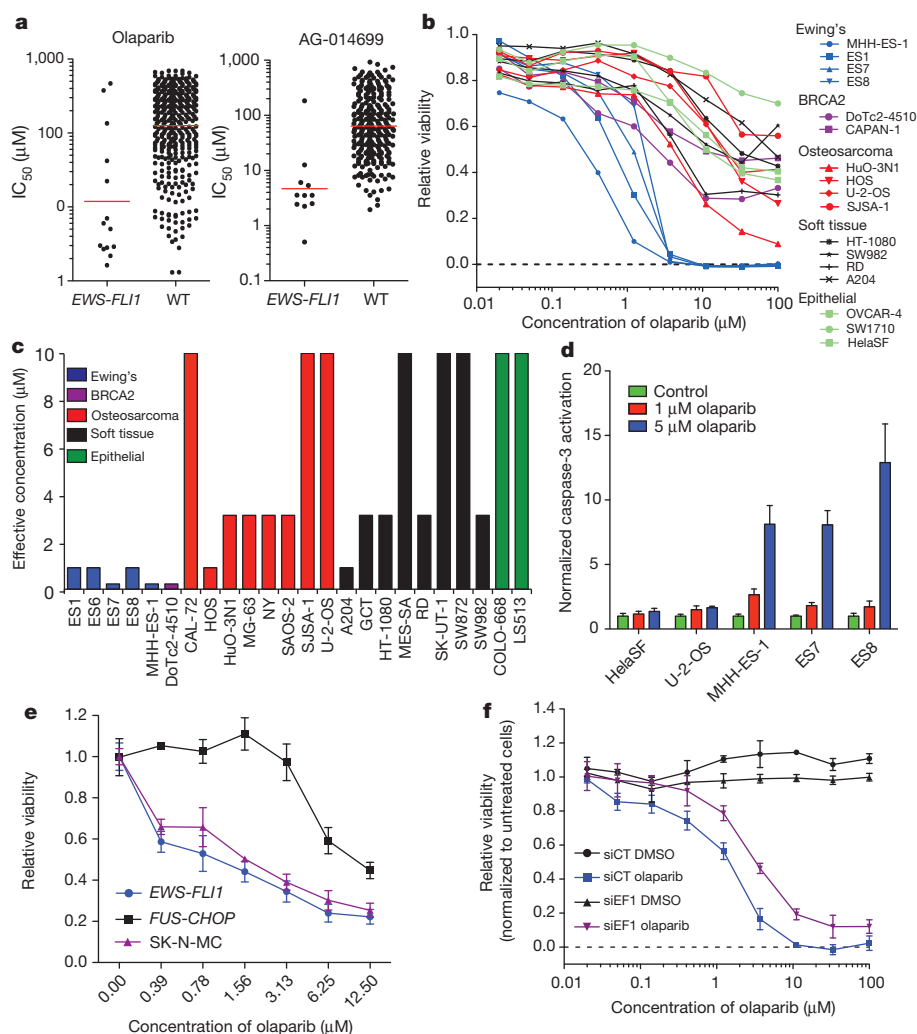


Figure 4 | Ewing's sarcoma cell lines are sensitive to PARP inhibition. **a**, The IC_{50} values of wild-type and *EWS-FLI1*-translocation-positive cell lines to olaparib and AG-014699. **b**, Dose-response curves to olaparib after 6 days of constant drug exposure. Cell lines are classified according to tissue subtype. **c**, Colony formation assays were performed for 7–21 days over a range of olaparib concentrations (0.1, 0.32, 1, 3.2 or 10 μM) and the concentration at which the number of colonies is reduced >90% for each cell line is indicated. **d**, Olaparib induced apoptosis in Ewing's sarcoma cell lines after 72 h treatment. **e**, Sensitivity to olaparib of *EWS-FLI1* and *FUS-CHOP* transformed mouse mesenchymal cells compared to the SK-N-MC cell line (which harbours the *EWS-FLI1* fusion). **f**, Sensitivity to olaparib of A673 cells transiently transfected with (siEF1) and without (siCT) *EWS-FLI1*-specific siRNA. All error bars are s.d. from triplicate measurements except for **b** where error bars have been removed for clarity.

To assess whether the sensitivity to PARP inhibitors is due to the presence of the *EWS-FLI1* rearrangement or intrinsic to the mesenchymal precursor cell type from which Ewing's sarcoma arises, we compared the sensitivity to olaparib of mouse mesenchymal cells transformed with either *EWS-FLI1* or the related liposarcoma-associated translocation *FUS-CHOP* (also known as *DDIT3*)^{27,28}. *EWS-FLI1*-transformed cells showed sensitivity comparable to human Ewing's sarcoma cells, whereas the *FUS-CHOP*-transformed cells were relatively resistant (Fig. 4e and Supplementary Fig. 18). Moreover, expression of *EWS-FLI1* in NIH3T3 cells conferred increased sensitivity to olaparib (Supplementary Fig. 19), whereas olaparib sensitivity was partially reverted by transiently depleting *EWS-FLI1* from Ewing's sarcoma cells (Fig. 4f and Supplementary Fig. 20). Higher *FLI1* expression was also correlated with sensitivity to olaparib even when considering only non-Ewing's sarcoma cell lines ($r = -0.32$ between IC_{50} and *FLI1* expression, $n = 413$, $P = 1.68 \times 10^{-11}$), suggesting that the sensitivity to olaparib of Ewing's sarcoma lines might be related to *EWS-FLI1* transcriptional activity.

Mutations in *BRCA1* or *BRCA2* are not present in these Ewing's sarcoma lines (Supplementary Data 1), and we have observed no evidence to indicate that the DNA damage response is defective in Ewing's sarcoma cells (data not shown). However, for reasons that are currently unclear, the *EWS-FLI1* translocation was associated with sensitivity to cytotoxic drugs, including DNA damaging agents such as camptothecin ($P < 1 \times 10^{-5}$), cisplatin ($P < 1 \times 10^{-4}$) and mitomycin-C ($P < 0.001$) (Supplementary Figs 21 and 22). Together with the report of olaparib sensitivity in prostate cancer cell lines

bearing the translocation between *ERG*, which like *FLI1* is a member of the *ETS* gene family, and *TMPRSS2*²⁹, our data support the potential utility of *ETS* gene fusions as biomarkers of sensitivity to PARP inhibitors. Notably, however, unlike the effect reported in prostate cancer, we observe potent cell death with olaparib treatment alone in Ewing's sarcoma cells. These observations raise the possibility that PARP inhibitors would be useful in the treatment of Ewing's sarcoma, a tumour of children and young adults with a 15% five-year survival rate in patients with metastatic disease or relapse after chemotherapy³⁰.

Discussion

High-throughput cancer cell line screening for drug-sensitivity patterns provides a strategy to identify appropriate cancer subtypes and biomarkers that may guide the early-phase clinical trials of multiple novel compounds under development. The validity of this approach is supported by its effective identification of drug-genotype associations that have already been established clinically, and it sets the stage for clinical testing of novel therapeutic biomarkers, such as the association between the *EWS-FLI1* translocation in Ewing's sarcoma cells and sensitivity to PARP inhibitors. The Supplementary Data accompanying this report, as well as the ongoing public web resource from future screenings (Genomics of Drug Sensitivity in Cancer; <http://www.cancerRxgene.org>), will hopefully enhance the discovery and validation of additional predictive cancer biomarkers.

Although the large number of cell lines screened facilitates representation of rare cancer genotypes and mitigates against the effects of individual samples, the data set presented here is limited both by the number

of available genotypes, as well as the number of targets interrogated by currently available drugs. Despite the apparent utility of using tumour-derived cell lines grown in two-dimensional culture, it is likely that experimental models that better mimic the tumour environment would in some instances further improve our understanding of drug sensitivity and provide additional insights. Nonetheless, we can discern an initial landscape of drug sensitivity patterns across a broad set of different cancer types and genomic backgrounds.

The identification of 'outliers' that are exquisitely sensitive to a drug as a result of a specific genetic abnormality within a targeted pathway remains the most compelling vision for targeted cancer therapies. *BCR-ABL*-positive CML, *BRAF*-mutant melanoma and *EGFR*-mutant-positive lung cancer and drugs that target the protein products of these genes are now well-established associations. The observation of PARP-inhibitor sensitivity by *EWS-FLI1*-positive Ewing's sarcoma cell lines points to the likelihood of new potent gene–drug associations as novel chemical and genomic space are explored. Even in the absence of outlier effects, pharmacogenomic profiling reveals a wealth of biomarkers that may prove useful for patient stratification. Although further work is needed to assess their potential clinical utility, in some instances these biomarkers may help explain heterogeneity in clinical responsiveness even among preselected patient populations.

This work, as well as an accompanying report³¹, provides a systematic and extensive view of the genomics underlying the sensitivity of human cancer cell lines to the diverse array of cancer drugs currently in use or under development. The emergent picture is of a complex network of biological factors that affect sensitivity to the majority of cancer drugs. This underscores both the challenge of identifying preselected patient populations for targeted therapies, as well as the opportunity to improve existing therapies and find new therapeutic avenues by identifying more predictive biomarkers.

METHODS SUMMARY

Cells were treated with nine concentrations (twofold dilutions) of drug for 72 h before measuring cell number relative to controls. A MANOVA was used to examine how drug IC₅₀ and slope values associate with tissue type, the mutation status of 64 cancer genes (including gene amplifications and homozygous deletions), rearrangements and MSI. The elastic net regression used the same genomic data sets as the MANOVA and also incorporated additional copy number data from a total of 426 cancer genes, transcriptional profiles and tissue types to identify features associated with drug response as measured by cell line IC₅₀.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 July 2011; accepted 2 March 2012.

1. Druker, B. J. *et al.* Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N. Engl. J. Med.* **355**, 2408–2417 (2006).
2. Kwak, E. L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* **363**, 1693–1703 (2010).
3. Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with *BRAF* V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011).
4. McDermott, U. & Settleman, J. Personalized cancer therapy with selective kinase inhibitors: an emerging paradigm in medical oncology. *J. Clin. Oncol.* **27**, 5650–5659 (2009).
5. Shoemaker, R. H. *et al.* Development of human tumor cell line panels for use in disease-oriented drug screening. *Prog. Clin. Biol. Res.* **276**, 265–286 (1988).
6. Weinstein, J. N. *et al.* An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343–349 (1997).
7. McDermott, U. *et al.* Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling. *Proc. Natl Acad. Sci. USA* **104**, 19936–19941 (2007).
8. Suwaki, N. *et al.* A HIF-regulated VHL-PTP1B-Src signaling axis identifies a therapeutic target in renal cell carcinoma. *Sci. Transl. Med.* **3**, 85ra47 (2011).
9. Deng, L. *et al.* Rho-kinase inhibitor, fasudil, suppresses glioblastoma cell line progression *in vitro* and *in vivo*. *Cancer Biol. Ther.* **9**, 875–884 (2010).
10. Weber, D. M. *et al.* Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. *N. Engl. J. Med.* **357**, 2133–2142 (2007).
11. Konecny, G. E. *et al.* Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res.* **66**, 1630–1639 (2006).

12. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).
13. O'Farrell, A. M. *et al.* SU11248 is a novel FLT3 tyrosine kinase inhibitor with potent activity *in vitro* and *in vivo*. *Blood* **101**, 3597–3605 (2003).
14. Kunii, K. *et al.* *FGFR2*-amplified gastric cancer cell lines require *FGFR2* and *ErbB3* signaling for growth and survival. *Cancer Res.* **68**, 2340–2348 (2008).
15. Byron, S. A. *et al.* Inhibition of activated fibroblast growth factor receptor 2 in endometrial cancer cells induces cell death despite PTEN abrogation. *Cancer Res.* **68**, 6902–6907 (2008).
16. Vassilev, L. T. *et al.* *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–848 (2004).
17. Konecny, G. E. *et al.* Expression of p16 and retinoblastoma determines response to CDK4/6 inhibition in ovarian cancer. *Clin. Cancer Res.* **17**, 1591–1602 (2011).
18. Boisvert-Adamo, K., Longmate, W., Abel, E. V. & Apelin, A. E. Mcl-1 is required for melanoma cell resistance to anoikis. *Mol. Cancer Res.* **7**, 549–556 (2009).
19. Brown, C. J., Lain, S., Verma, C. S., Fersht, A. R. & Lane, D. P. Awakening guardian angels: drugging the p53 pathway. *Nature Rev. Cancer* **9**, 862–873 (2009).
20. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* **67**, 301–320 (2005).
21. Hanafusa, H., Torii, S., Yasunaga, T. & Nishida, E. Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nature Cell Biol.* **4**, 850–858 (2002).
22. Patterson, K. I., Brummer, T., O'Brien, P. M. & Daly, R. J. Dual-specificity phosphatases: critical regulators with diverse cellular targets. *Biochem. J.* **418**, 475–489 (2009).
23. Dry, J. R. *et al.* Transcriptional pathway signatures predict MEK addiction and response to selumetinib (AZD6244). *Cancer Res.* **70**, 2264–2273 (2010).
24. Kelland, L. R., Sharp, S. Y., Rogers, P. M., Myers, T. G. & Workman, P. DT-Diaphorase expression and tumor cell sensitivity to 17-allylamino, 17-demethoxygeldanamycin, an inhibitor of heat shock protein 90. *J. Natl Cancer Inst.* **91**, 1940–1949 (1999).
25. Fong, P. C. *et al.* Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N. Engl. J. Med.* **361**, 123–134 (2009).
26. McCabe, N. *et al.* BRCA2-deficient CAPAN-1 cells are extremely sensitive to the inhibition of poly(ADP-ribose) polymerase: an issue of potency. *Cancer Biol. Ther.* **4**, 934–936 (2005).
27. Riggi, N. *et al.* Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res.* **65**, 11459–11468 (2005).
28. Riggi, N. *et al.* Expression of the FUS-CHOP fusion protein in primary mesenchymal progenitor cells gives rise to a model of myxoid liposarcoma. *Cancer Res.* **66**, 7016–7023 (2006).
29. Brenner, J. C. *et al.* Mechanistic rationale for inhibition of poly(ADP-ribose) polymerase in ETS gene fusion-positive prostate cancer. *Cancer Cell* **19**, 664–678 (2011).
30. Balamuth, N. J. & Womer, R. B. Ewing's sarcoma. *Lancet Oncol.* **11**, 184–192 (2010).
31. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* <http://dx.doi.org/10.1038/nature11003> (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Lo Grasso of the Scripps Research Institute for providing the inhibitor JNK9L. This work was supported by a grant from the Wellcome Trust (086357; M.R.S., P.A.F., J.S., D.A.H.) and by grants from the National Institutes of Health (P41GM079575-02 to N.S.G. and 1U54HG006097-01 to N.S.G. and D.A.H.). S.R. is supported by a Physician-Scientist Early Career Award from the Howard Hughes Medical Institute. U.M. is supported by a Cancer Research UK Clinician Scientist Fellowship.

Author Contributions M.J.G., C.H.B., U.M. and S.V.S. supervised data collection. M.J.G., C.H.B., U.M. and S.R. supervised data analysis. C.D.G. and K.W.L. conceived and wrote the curve-fitting algorithm and performed the MANOVA; E.J.E. and S.R. performed elastic net analysis and analysed the data. P.G., I.R.T. and J.S. developed and managed screening databases with assistance from A.B. and W.Y.; S.J.H. performed most of the Ewing's sarcoma related studies with contributions from D.S., A.D., X.L., F.K. and L.C., with I.S. and O.D. providing critical reagents; R.J.M., A.T.T., J.A.S., S.B., S.R.L., K.L., A.M.-D., J.L.B., X.M., T.M., H.T., L.R., F.J. and P.O'B. performed cell line screening experiments. S.P. performed *MCL1* siRNA experiments. Q.L.W.Z., T.Z., W.H., X.D., H.G.C. and J.W.C. synthesized screening compounds, and N.S.G. provided guidance on their selection and use; F.I. and J.S.-R. performed compound activity clustering; G.R.B. and H.D. performed cell line genotyping and genetic analysis; J.A.E. and J.B. provided guidance regarding clinical relevance of the work; M.J.G. and C.H.B. wrote the manuscript with major contributions from S.R., S.J.H. and U.M.; M.R.S., D.A.H., J.Se. and P.A.F. conceived the study, analysed the data and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to U.M. (um1@sanger.ac.uk) or C.B. (cbenes@partners.org).

METHODS

Cell lines. All cell lines were sourced from commercial vendors. Cells were grown in RPMI or DMEM/F12 medium supplemented with 5% FBS and penicillin/streptavidin, and maintained at 37 °C in a humidified atmosphere at 5% CO₂. Cell lines were propagated in these two media to minimize the potential effect of varying the media on sensitivity to therapeutic compounds in our assay, and to facilitate high-throughput screening. To exclude cross-contaminated or synonymous lines, a panel of 92 SNPs was profiled for each cell line (Sequenom) and a pair-wise comparison score was calculated. In addition, we performed short tandem repeat (STR) analysis (AmpFISTR Identifier, Applied Biosystems) and matched this to an existing STR profile generated by the providing repository. More information on the cell lines screened, including their SNP and STR profiles is available on the Genomics of Drug Sensitivity in Cancer website (<http://www.cancerRxgene.org>).

Screening drugs. Compounds were from academic collaborators or commercial vendors. We have provided a full description for each compound including its name, source, PubChem and/or ChEMBL accessions, screening concentration as well as therapeutically relevant molecular target(s) (Supplementary Data 2). Compounds were stored as 10 mM aliquots at –80 °C, and were subjected to a maximum of five freeze–thaw cycles. The range of concentrations selected for each compound was based on *in vitro* data of concentrations inhibiting relevant kinase activity and cell viability.

Cell viability assays. Cells were seeded in either 96-well or 384-well microplates in medium supplemented with 5% FBS and penicillin/streptavidin. The optimal cell number for each cell line was determined to ensure that each was in growth phase at the end of the assay (~70% confluency). Adherent cell lines were plated 1 day before treatment with a 9-point twofold dilution series of each compound using liquid handling robotics, and assayed at a 72-h time point. Cells were fixed in 4% formaldehyde for 30 min and then stained with 1 µM of the fluorescent nucleic acid stain Syto60 (Invitrogen) for 1 h. Suspension cell lines were treated with compound immediately following plating, incubated for 72 h, and then stained with 55 µg ml^{–1} resazurin (Sigma) prepared in glutathione-free media for 4 h. Quantification of fluorescent signal intensity was performed using a fluorescent plate reader at excitation and emission wavelengths of 630/695 nm for Syto60, and 535/595 nm for resazurin. All screening plates were subjected to stringent quality control measures and a Z-factor score comparing negative and positive control wells was calculated across all screening plates (median = 0.70, upper quartile = 0.86, lower quartile = 0.47, $n = 4,857$ plates). Drug screening was performed at two sites using matched cell line collections (the site of drug screening and plate format for each IC₅₀ are described in Supplementary Data 2 and 10). As a control the drug camptothecin (drugs number 1003 and 195) was screened at both sites. The IC₅₀ values were highly correlated ($r^2 = 0.3244$, slope = 1.030, $n = 252$ cell lines) and these drugs were nearest neighbours in our cluster analysis of drugs (Supplementary Table 1). Measurements of cell viability during 6-day assays using threefold dilution series of olaparib were performed in 96-well plates using Cell Titer Blue (Promega) according to the manufacturer's instructions. A673 and NIH3T3 cells were also exposed to a threefold dilution series of olaparib and cell viability was measured after 72 h of drug exposure using Cell Titer 96 Aqueous One Solution Cell (Promega) according to the manufacturer's instructions. Measurements of cellular apoptosis were performed using Apo-ONE caspase assay (Promega) following manufacturer's instructions. The A673 Ewing's sarcoma cell line was transiently transfected with allstars non-targeting siRNA control (Qiagen; siCT) or an siRNA targeting the *EWS-FLI1* translocation³² (siEF1, 5'-GGCAGCAGAACCCUUCUACG) and treated with olaparib immediately after siRNA transfection. Transient knock-down of *MCL1* in HT-144 and A549 cells was performed using the following siRNA sequences: oligonucleotide 1, 5'-CTGGTTTGGCATATCTAATAA; oligonucleotide 2, 5'-CCCGCCGAATTCATTAATTTA; oligonucleotide 3, 5'-AAGGGTTAGGACCAACTACAA; oligonucleotide 4, 5'-CCCTAGCAACC TAGCCAGAAA) and controls were mock transfected.

Colony formation assays. Cells were plated at low density into 35-mm cell-culture plates and the following day treated with the indicated drug concentration or vehicle control (DMSO). The medium was changed and cells re-drugged every 3–4 days. When sufficient colonies were visible, typically after 7–21 days, cells were washed once in PBS before fixing in ice-cold methanol for 30 min while shaking. Methanol was aspirated and Giemsa stain added at a dilution of 1:20 overnight while shaking. The following day cells were rinsed in distilled water and air dried.

Genomic and transcriptional characterization of cancer cell lines. A total of 64 of the most frequently mutated cancer genes were sequenced to base-pair resolution across all coding exons for each gene by capillary sequencing in our panel of human cancer cell lines, which formed the basis for the cell lines chosen for this drug screen. The presence of seven of the most commonly rearranged cancer

genes (for example, *BCR-ABL*, *MLL-AFF1* and *EWS-FLI1*) was determined across the drug screen cell line panel by the design of breakpoint-specific sequence primers that enabled the detection of the rearrangement following capillary sequencing. Analysis of MSI was carried out according to the guidelines set down by The International Workshop on Microsatellite Instability and RER Phenotypes in Cancer Detection and Familial Predisposition³³. Samples were screened using the markers BAT25, BAT26, D5S346, D2S123 and D17S250 and were characterized as MSI if two or more markers showed instability. Total integral copy number values across the footprints of the cancer genes were determined from Affymetrix SNP6.0 microarray data using the PICNIC algorithm to predict copy number segments in each of the cell lines³⁴. For a gene to be classified as amplified, the entire coding sequence must be contained in one contiguous segment defined by PICNIC, and have a total copy number of eight or more. A copy number of eight was chosen to ensure that the majority of segments above this threshold consisted of focal amplifications more likely to have been selectively amplified. Deletions must occur within a single contiguous segment with copy number zero. For gene expression analysis, RNA was extracted from each cell line using a standard Trizol protocol and hybridized to the HT-HGU133A Affymetrix whole genome array. Normalized gene expression intensities were generated using the Robust Multi-Array Average (RMA) algorithm³⁵. The genomic data used for this analysis are provided in the Supplementary Data and transcriptional data are available from ArrayExpress (accession number E-MTAB-783). A complete description of the characterization of our cancer cell line collection is also available from the Cancer Genome Project webpages (<http://www.sanger.ac.uk/genetics/CGP/CellLines/>).

Curve fitting of drug sensitivity data. Dose–response curves were fitted to the fluorescence signal intensities. We required a method that would first allow us to model the heteroscedasticity in the luminescence data, and second allow us to incorporate prior knowledge of response, especially at drug concentrations at which the data are less informative. A bespoke Bayesian sigmoid model was thus implemented to facilitate this, yielding a full description of the uncertainties in the data, and allowing reasonable interpretation of predicted response at concentrations outside the tested range. Response curves were fitted to the fluorescence signal intensities using a Bayesian sigmoid model. Drug–response data consisted of 16 (96-well format) or 42 (384-well) drug-free positive controls, 8 (96-well) or 32 (384-well) negative (no cells) controls and drug–response points for nine half-fold concentrations. Technical replicate intensity responses were averaged. Generalized sigmoidal response curves are fitted as follows.

$$\text{Intensity } x_{lc} \text{ is assumed to have the mean value } x_{lc} = I_{\min} + \frac{I_{\max} - I_{\min}}{(1 + e^{\beta(lc - z)})^f}.$$

Parameters I_{\max} and I_{\min} are the mean intensities of the positive and negative controls, α and β are scale and gradient responses, f is a shape parameter, lc denotes the log concentration. We assume that the intensity x_{lc} has variance $\text{Var}(x_{lc}) = BE(x_{lc})$, where B represents a noise parameter. We assume that x_{lc} has a gamma distribution:

$$P(x_{lc}) = \frac{B^{E(x_{lc})/B}}{\Gamma(E(x_{lc})/B)} x_{lc}^{E(x_{lc})/B - 1} e^{-Bx_{lc}}$$

Positive and negative controls x_{\max} and x_{\min} are also gamma distributed:

$$P(x_{\max}) = \frac{B^{I_{\max}/B}}{\Gamma(I_{\max}/B)} x_{\max}^{I_{\max}/B - 1} e^{-Bx_{\max}}$$

$$P(x_{\min}) = \frac{B^{I_{\min}/B}}{\Gamma(I_{\min}/B)} x_{\min}^{I_{\min}/B - 1} e^{-Bx_{\min}}$$

The concentration giving percentage response p is given by:

$$IC_p = \alpha + \frac{1}{\beta} \log\left(\left(\frac{100}{100-p}\right)^{\frac{1}{f}} - 1\right)$$

We use Markov Chain Monte Carlo simulations to obtain mean posterior parameter estimates. Parameters were initialized from maximum likelihood estimates and 100,000 iterations obtained. The final 10,000 values were used for subsequent inference. The IC₅₀ has a normal prior with 95% probability mass covering a range from 1,000-fold below the minimum concentration tested to 1,000-fold higher than the maximum tested concentration. We assume uninformative priors on the remaining parameters. Response curves are plotted using mean posterior values of IC_p for p ranging between 0% and 100%. Confidence intervals for IC_p are obtained from the associated posterior. Results were manually curated as part of quality control.

MANOVA. A fixed effects MANOVA was used to correlate response with genomics. An $n \times 2$ dose–response matrix consisting of IC₅₀ and slope parameter β for n cell lines was constructed for each drug. A linear (no interaction terms)

model explained these observables with factors including tissue type, the mutation status of cancer genes, chromosomal re-arrangements, and MSI status. Size effects and significances were obtained. A gene was defined as mutated if it fulfilled any of these criteria: a coding sequence variant in the cancer gene, a total copy number of 0 (homozygous deletion) or more than 7 (amplification). Only those genes with >1 mutated cell lines in the panel were included for analysis (65 cancer genes and 3 chromosomal re-arrangements in total). The effect measures the relative difference in the mean IC₅₀ from the wild-type to mutant group (for example, an effect of 0.1 or 10 indicates a ~10-fold decrease or increase in drug concentration, respectively). A Benjamini–Hochberg multiple testing correction threshold with false discovery rate of 20% was used to identify a candidate list of significant associations for the purposes of this paper.

Elastic net analysis. We take the previously described approach of the elastic net²⁰, a multivariate variable selection technique with a penalization approach. Genomic data including mutation status of 64 cancer genes, rearrangements, continuous copy number data from 426 genes causally implicated in cancer (<http://www.sanger.ac.uk/genetics/CGP/Census/>), and genome-wide transcriptional profiles, as well as tissue type, were used as input variables (Supplementary Data 1 and 11). The elastic net was used to select which of these features were associated with drug response as measured by IC₅₀ across the cell line panel.

Let X be a $n \times p$ matrix of input features (where p is the number of features and n is the number of cell lines) and y be a vector of drug sensitivities of length n . For any non-negative λ_1 and λ_2

$$L(\lambda_1, \lambda_2, \beta) = |y - X\beta|^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \quad (1)$$

Let $\hat{\beta}$ be the naive elastic net estimator. Then $\hat{\beta} = \arg \min_{\beta} \{L(\lambda_1, \lambda_2, \beta)\}$. A scaling factor of $(1 + \lambda_2)$ is added to the naive elastic net to prevent double shrinking.

$$\hat{\beta}(\text{elasticnet}) = (1 + \lambda_2) \hat{\beta}(\text{naiveelasticnet}) \quad (2)$$

To determine the optimal λ_1 and λ_2 , we let $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$. Then $(1 - \alpha) \sum_{j=1}^p \beta_j^2 + \alpha \sum_{j=1}^p |\beta_j|$ is the elastic net penalty. Tenfold cross validation is performed to optimize λ_1 and λ_2 in equation (2), denoted as $\hat{\lambda}_1$ and $\hat{\lambda}_2$, respectively. To find the variables that are associated with drug response, $\hat{\lambda}_1$, $\hat{\lambda}_2$, X and y are inputted into equation (2) to solve for vector β . The variables with non-zero β values are determined to be features associated with drug response.

This procedure was repeated 100 times for each drug to assess the stability of the features when applying the tenfold cross validation procedure. For each of the 100 runs, a feature list was built for the drug comprised of genes, transcripts, and tissues with weights assigned to each. The final signature of markers for a drug consisted of all features that appear in any of the 100 runs along with the statistics on the frequency that the feature appears and the average weight given to that feature over the 100 runs.

The weights (β) were used to assess effect sizes of features in a drug's marker signature. The effect size of a feature was calculated by multiplying the feature's weight by its standard deviation across the cell line panel. The effect size is therefore a normalization of the feature's weight to account for the different scales used to measure the different genomic features. Features with higher stability of correlation in cross-validation (f) were considered of the highest confidence of truly being associated with drug response. The most significant features associated with drug response are those with both large frequency and effect size. Highly significant associations are defined as those with an effect size that was ± 2 s.d. from the mean, $-2.95 > e > 2.79$ and a frequency (f) of 2 s.d. above the mean, $f > 0.76$.

Drug clustering based on IC₅₀ values. The data set was filtered to remove cell lines for which less than 50% of drugs in the panel were tested, resulting in 400 cell lines remaining. Missing data points were inputted by using a nearest-neighbour approach based on Pearson correlation (PC) scores between IC₅₀ values. We also applied a step to reduce effects inherently present in the data due to the integration of data that were obtained from two different screening sites and which may be amplified by the nearest-neighbour missing point imputation. To do this, we performed an ANOVA for each cell line to assess the difference in the average IC₅₀ values for drugs at the two screening sites. Cell lines with a P value < 0.01 according to this test were filtered out. Of the original data set, 265 cell lines (41.47%) were used for cluster analysis.

As done for missing point imputation, we measured the extent of similarity among drug sensitivity/resistance profiles with PC scores pair-wisely by considering for each drug its pattern of IC₅₀ values across the cell lines (log values were considered). Then we used these similarity scores as input to the affinity-propagation clustering (APC) algorithm in a recursive manner³⁶. The APC algorithm is based on a 'passing-message-between-datapoints' strategy, it requires as input a pair-wise similarity matrix and gives as the output a set of disjointed clusters. It also indicates, for each of the computed clusters, a datapoint called the 'cluster exemplar' (that is, the 'cluster centroid'): the datapoint that best interpolates all the other datapoints in its cluster. The APC algorithm requirement consists in a square matrix containing similarity scores between each pair of datapoints and a set of probabilities, one for each datapoint to be elected as the exemplar of its clusters. Implicitly, these probabilities determine the number of clusters to be computed. If they are uniformly distributed then the APC algorithm from the data automatically determines the number of clusters. In the first step of our clustering strategy, we applied the APC algorithm to the whole set of drugs in the screening by assuming the probability of each drug to be elected as a cluster exemplar uniformly distributed (so the number of clusters was automatically determined by the APC algorithm). Particularly, we set the input parameter from which this probability is computed equal to the median value of all the pair-wise similarities. The algorithm provided 22 clusters. In each of these clusters, a drug was indicated as the cluster exemplar. The intra-cluster similarity score (odds ratio) of a community C containing n drugs is computed as the average similarity (that is, correlation between IC₅₀ patterns) between all the possible couple of drugs belonging to C (total average correlation) divided by the expected average similarity between all the possible couple of drugs in a randomly selected set of n drugs. We derived a network representation of the clustered data by considering each drug as a network node and by adding a link between a couple of nodes corresponding to drugs in the same cluster. This generated the 'network communities' indicated with different colours in Supplementary Fig. 3. We then, recursively, clustered again the cluster exemplars with the APC algorithm, to obtain 'communities of communities' (or 'rich clubs') and added links to the network as explained above. This procedure was recursively re-applied over cluster exemplars until convergence (no datapoints were clustered together) ending up into a hierarchical network depicted in Supplementary Fig. 3 and as described in ref. 37. The network visualization in Supplementary Fig. 3 has been obtained by using Cytoscape³⁸.

The inter-cluster similarity score for each couple of communities A and B was computed as follows:

$$\rho(A, B) = \frac{\sum_{X \in A} \sum_{Y \in B} \rho_{X,Y}}{|A||B|}$$

Where $\rho_{X,Y}$ is the Pearson correlation coefficient between the patterns of IC₅₀ values of drug X and drug Y .

Empirical P values were computed for each of these scores by permutation test: for a number $n = 10,000$ of trials and each $\rho(A, B)$, we sampled two random sets of drugs A^* and B^* (containing $|A|$ drugs and $|B|$ drugs, respectively) from the whole set of 131 drugs and we computed $\rho(A^*, B^*)$. Then we computed the empirical P value for a given $\rho(A, B) \geq 0$ (respectively ≤ 0) as the number of times that $\rho(A^*, B^*)$ was greater (resp. less) than or equal to $\rho(A, B)$ across the n permutation trials, divided by n .

32. Prieur, A., Tirode, F., Cohen, P. & Delattre, O. EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3. *Mol. Cell. Biol.* **24**, 7275–7283 (2004).
33. Boland, C. R. et al. A National Cancer Institute Workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* **58**, 5248–5257 (1998).
34. Greenman, C. D. et al. PICNIC: an algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics* **11**, 164–175 (2010).
35. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
36. Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972–976 (2007).
37. Iorio, F. et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl Acad. Sci. USA* **107**, 14621–14626 (2010).
38. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

Observation of a roton collective mode in a two-dimensional Fermi liquid

Henri Godfrin¹, Matthias Meschke^{1,2}, Hans-Jochen Lauter^{3,4}, Ahmad Sultan¹, Helga M. Böhm⁵, Eckhard Krotscheck^{5,6} & Martin Panholzer⁵

Understanding the dynamics of correlated many-body quantum systems is a challenge for modern physics. Owing to the simplicity of their Hamiltonians, ⁴He (bosons) and ³He (fermions) have served as model systems for strongly interacting quantum fluids, with substantial efforts devoted to their understanding. An important milestone was the direct observation of the collective phonon–roton mode in liquid ⁴He by neutron scattering, verifying Landau’s prediction¹ and his fruitful concept of elementary excitations. In a Fermi system, collective density fluctuations (known as ‘zero-sound’ in ³He, and ‘plasmons’ in charged systems) and incoherent particle–hole excitations are observed. At small wavevectors and energies, both types of excitation are described by Landau’s theory of Fermi liquids^{2,3}. At higher wavevectors, the collective mode enters the particle–hole band, where it is strongly damped. The dynamics of Fermi liquids at high wavevectors was thus believed to be essentially incoherent. Here we report inelastic neutron scattering measurements of a monolayer of liquid ³He, observing a roton-like excitation. We find that the collective density mode reappears as a well defined excitation at momentum transfers larger than twice the Fermi momentum. We thus observe unexpected collective behaviour of a Fermi many-body system in the regime beyond the scope of Landau’s theory. A satisfactory interpretation of the measured spectra is obtained using a dynamic many-body theory⁴.

Quantum many-body systems are ubiquitous in nature; the identification of their ground state and the description of their elementary excitations is a cornerstone of modern physics. Nuclei, metals, semiconductors and neutron-star matter are all examples of quantum

fluids. Their properties depend on the quantum statistics obeyed by their constituent particles (electrons, nucleons, atoms), leading to their classification as Bose or Fermi systems. Weakly interacting Bose and Fermi systems are well understood, but extending this understanding to their strongly interacting analogues has not been straightforward, with much work having been devoted to correlated quantum systems^{1–6}.

Our work centres on the interplay between statistics and interactions in quantum many-body systems—specifically, in liquid ⁴He and ³He, the canonical Bose and Fermi quantum fluids. Whereas liquid ⁴He becomes superfluid at 2.17 K, liquid ³He remains a normal Fermi liquid down to millikelvin temperatures, where Cooper pairs form and condense into several superfluid phases. Clearly, Bose and Fermi liquids behave differently, and are thus expected to sustain very different excitations.

Landau¹ described the elementary excitations of a Bose fluid. Their dispersion relation (Fig. 1) shows a sharp, linear ‘phonon’ mode, which evolves continuously as a function of the wavevector, displaying a pronounced ‘roton’ minimum^{1,6–9}. The excitations remain well defined even at atomic wavevectors and at relatively high temperatures. Modern many-body theories have proven successful in describing the dynamics of Bose fluids at different densities and dimensionalities: bulk, films or droplets^{10–12}.

The low-lying elementary excitations of liquid ³He (Fig. 2) are incoherent particle–hole excitations, as well as collective density and

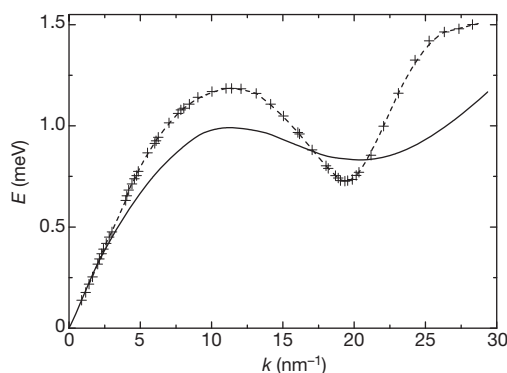


Figure 1 | Elementary excitations of superfluid ⁴He. The solid line is the dispersion relation predicted by Landau¹; crosses correspond to the excitation energy as a function of wavevector, determined by neutron scattering⁸. At low wavevectors, the dispersion relation is linear, and the excitations are quantised sound waves (phonons). At higher wavevectors, the spectrum evolves continuously, displaying a maximum and then a characteristic minimum. The corresponding excitations are called, respectively, maxons and rotons; the latter play an essential role in the thermodynamic properties of superfluid ⁴He.

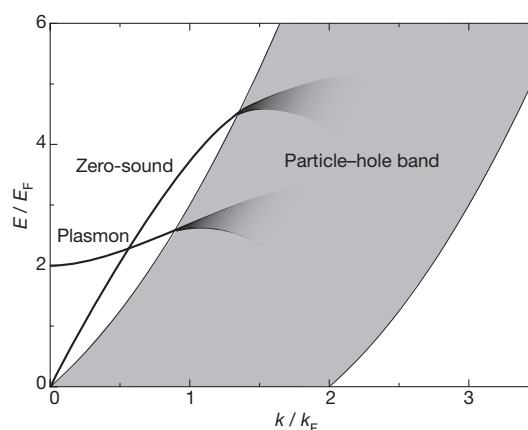


Figure 2 | Schematic picture of the elementary excitations of a Fermi liquid. The broad shaded area corresponds to the particle–hole band, that is, to the range of excitation energy as a function of wavevector accessible by promoting a particle occupying a state inside the Fermi surface, to an empty state outside it. In addition, an interacting Fermi system displays collective density modes, called ‘plasmons’ in charged systems, and ‘zero-sound’ in neutral systems. With increasing wavevector, the collective modes enter the particle–hole band, where they decay (Landau damping) into incoherent particle–hole excitations.

¹Institut Néel, CNRS et Université Joseph Fourier, BP 166, 38042 Grenoble Cedex 9, France. ²Low Temperature Laboratory, Aalto University, PO Box 15100, 00076 Aalto, Finland. ³Institut Laue-Langevin, BP 156, 38042 Grenoble Cedex 9, France. ⁴Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, Tennessee 37831-6475, USA. ⁵Institute for Theoretical Physics, Johannes Kepler University, A-4040 Linz, Austria. ⁶Department of Physics, University at Buffalo SUNY, Buffalo, New York, 14260, USA.

spin-density modes, described by Landau's Fermi liquid theory^{2,3,8}. For intermediate energies, a qualitative description is provided by the random phase approximation (RPA)^{3,5,8}. In the RPA, particle-hole states are confined (Fig. 2) within the particle-hole band (PHB). The boundaries of the PHB for a non-interacting system (the Fermi 'gas') are $E_{\text{max,min}}/E_F = (k/k_F)^2 \pm 2(k/k_F)$, where $E_F = \hbar^2 k_F^2/2m$ is the Fermi energy, k the excitation wavevector, k_F the Fermi wavevector, and m the (bare) mass of a particle.

Landau's Fermi liquid theory postulates that an interacting system behaves like a Fermi gas with renormalized parameters. In particular, an effective mass m^* is assigned to the fermionic 'quasi-particles'. The theory describes well the low-temperature properties of bulk liquid ³He, where m^* , depending on the liquid pressure, varies from 3 to 6 times m . This picture does not apply at high wavevectors; it was theoretically shown^{13–15}, and experimentally verified¹⁶ in bulk liquid ³He, that the effective mass enhancement is confined to the vicinity of the Fermi surface. Therefore, the PHB is expected to be essentially that calculated for the non-interacting system, except for very low energies, where it is strongly depressed.

The density collective mode, called zero-sound, is described by Landau as an oscillation of the whole Fermi sphere^{2,3,8}. Unlike ordinary sound, its frequency is higher than the collision rate. First detected by ultrasonic techniques, it has been investigated in detail by neutron scattering^{16–18}. Zero-sound (Fig. 2) has a linear dispersion relation above the PHB, then a negative deviation at intermediate wavevectors, and finally enters the PHB, where this mode disappears by decay into particle-hole excitations (Landau damping).

The corresponding excitation in electron fluids is the plasmon. Apart from an energy gap at zero wavevector due to charge, the physics is the same³. In particular, the plasmon dispersion curve is observed to enter the PHB, and to disappear, as shown in Fig. 2.

In an elegant discussion of the dynamics of Fermi many-body systems, Pines⁶ states that the phonon-roton mode of liquid ⁴He and the zero-sound mode of liquid ³He have a common origin in strong interactions, rather than quantum statistics. Nozières⁹ argued that the physical origin of the roton minimum may be the incipient localization of the particles due to interactions.

In the present experiment, we determine the dynamic structure factor of a monolayer of liquid ³He, essentially at zero temperature. Two-dimensional Fermi liquids have been extensively investigated by thermodynamic techniques^{19–24}; we present here a direct investigation of their elementary excitations by neutron scattering. We observe a collective mode, which remains visible throughout the whole PHB, and re-emerges as a well defined mode at large wavevectors, as shown in Fig. 3.

We make the helium film at low temperature by the controlled adsorption of gas onto a substrate, a high-quality ZYX exfoliated graphite (surface area 60 m²) with large coherence length (190 nm) and low mosaic spread (10°), essential for obtaining neutron spectra with good resolution^{24,25}. The substrate was first pre-plated with a complete monolayer of ⁴He. This high-density solid provides a smoother adsorption potential than bare exfoliated graphite. A monolayer of liquid ³He is then deposited onto the ⁴He-plated substrate. The amount of ⁴He introduced into the cell was $V_4 = 28.59 \text{ cm}^3 \text{ STP}$ (volume of gas at standard temperature and pressure). This is sufficient to complete the first monolayer, considering the effect of the pressure of the partial second layer. The amount of ³He was $V_3 = 11.0 \text{ cm}^3 \text{ STP}$. Using a coverage scale developed earlier²⁴ (see Supplementary Information), we obtain a lower limit $A_c = 59.7 \text{ m}^2$ and an upper limit $A_l = 65.3 \text{ m}^2$ for the area available for the ³He layer adsorption; the liquid ³He density is determined to be $\rho_3 = 4.7 \pm 0.2 \text{ atoms nm}^{-2}$. At this areal density, the ³He effective mass at the Fermi surface^{19–23} is $m^*/m \approx 4$, similar to that of bulk liquid ³He at a pressure of 1 MPa. An aluminium sample cell confines the gas during the adsorption process, performed through a filling capillary. Measurements are made in a dilution refrigerator, at temperatures below 100 mK.

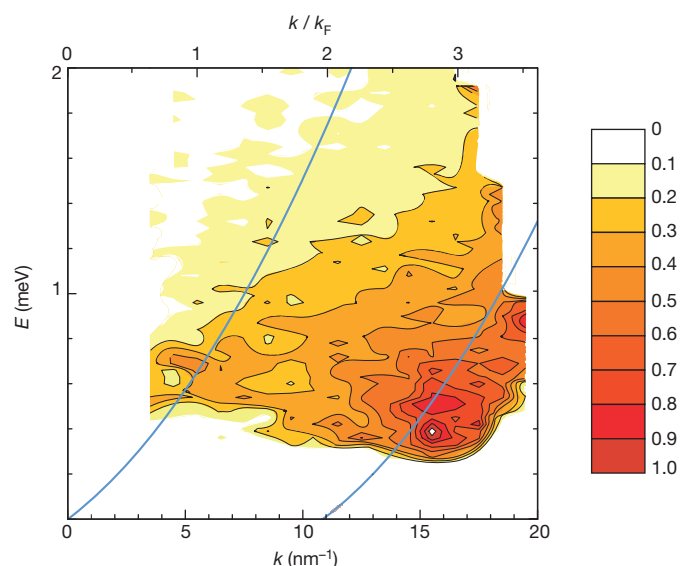


Figure 3 | Experimental dynamic structure factor. The structure factor $S(k, E)$ determined by inelastic neutron scattering for a monolayer of liquid ³He of areal density $4.7 \pm 0.2 \text{ atoms nm}^{-2}$, is shown as a function of the neutron momentum transfer $\hbar k$ and energy transfer E . The colour scale evolves from white to red (in arbitrary units proportional to neutron counts). White is also used in the lower part of the graph, where data cannot be exploited due to the large quasi-elastic background, and in the limits of low and high k determined by the angular range covered by the detectors. The blue lines show the limits of the particle-hole band of a Fermi gas with the bare ³He atomic mass. High-intensity regions indicate the existence of modes with wavevector k and energy E , broadened by the experimental resolution. The zero-sound collective mode, visible at $k \approx 5 \text{ nm}^{-1}$ and $E \approx 0.7 \text{ meV}$, is broadened inside the particle-hole band. It emerges beyond this band as a well defined mode, displaying a minimum as a function of energy at $k = 15.5 \text{ nm}^{-1}$ and $E = 0.4 \text{ meV}$. This fermionic collective mode closely resembles the phonon-maxon-roton dispersion relation of liquid ⁴He (see Fig. 1).

The experiments were performed at the Institut Laue-Langevin on the time-of-flight spectrometer IN6, using incident wavelengths of 0.512 and 0.41 nm. The measured dynamic structure factor $S(k, E)$ contains all the relevant information on the elementary excitations of a system; it gives the probability for creating an excitation with wavevector k and energy E . Figure 3 shows the main features revealed by our data. The zero-sound mode is seen at low wavevectors; given the limited experimental range, its definite identification requires theoretical support, to be presented later. It is found at energies well below those observed in bulk liquid ³He. The mode is broadened as it enters the PHB, and emerges beyond the limits of the PHB as an intense mode, displaying a minimum as a function of energy, and increasing rapidly beyond this minimum. The high-intensity region of $S(k, E)$ closely resembles the phonon-maxon-roton dispersion relation of liquid ⁴He. Significant intensity is present at low wavevectors above the PHB, demonstrating the existence of multi-pair excitations.

We now show that dynamic many-body effects play an essential role in explaining the observed position of the roton and the emergence of the collective mode beyond the PHB. Adopting the view that the physical mechanisms that determine the short-wavelength spectrum are the same in ⁴He and in ³He, we have developed the fermion generalization⁴ of the dynamic many-body theory of Jackson, Feenberg and Campbell^{10,12}. The boson theory has by now been brought to a level where a consistent description of the dynamics of ⁴He in the whole (k, E) plane is possible. The fermion version of the theory⁴ allows a calculation of the dynamics of strongly interacting systems in the language of a time-dependent Hartree-Fock theory⁷ with energy-dependent effective interactions. It supersedes the intuitive 'backflow' picture of Feynman⁷ and goes beyond the RPA by being applicable at atomic wavevectors.

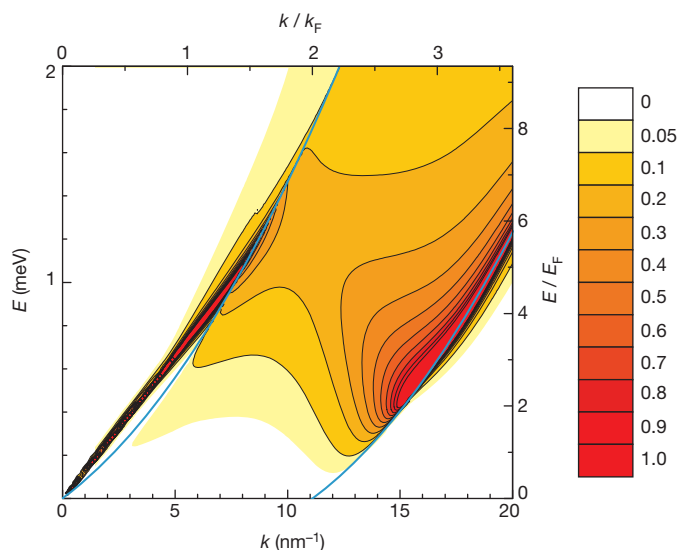


Figure 4 | Theoretical dynamic structure factor. The structure factor $S(k, E)$ calculated for a monolayer of liquid ^3He of areal density $4.9 \text{ atoms nm}^{-2}$, shown as a function of the wavevector k and energy E . At this areal density, the Fermi wavevector, k_F , is 5.55 nm^{-1} and the Fermi energy, E_F , is 0.213 meV . Blue lines as in Fig. 3. High-intensity regions indicate the existence of modes with wavevector k and energy E . The theoretical spectrum has been slightly broadened to make the sharp collective modes visible. The zero-sound collective mode, well defined at low wavevectors, enters the particle-hole continuum, is broadened, and finally emerges beyond the lower limit of the particle-hole band, displaying a minimum as a function of energy. A phonon-maxon-roton type of dispersion relation is clearly seen.

Figure 4 shows the results of our calculation of the dynamic structure factor at a density of $4.9 \text{ atoms nm}^{-2}$. We obtain good quantitative agreement with the experiment at a similar density, without any adjustable parameters. At low wavevectors, we observe a long-lived zero-sound collective mode, close to the PHB upper limit. The mode is broadened, but visible, within the PHB, and emerges from it as a well defined, intense excitation. A phonon-maxon-roton type of dispersion relation is clearly seen. Multi-pair excitations are present at low wavevectors above the PHB, causing a natural width of the phonon.

The contour plots of Figs 3 and 4 show a strong down-shift in the density of states compared to the RPA predictions. This is more clearly seen in the spectra at constant wavevector presented in Fig. 5. The experiments yield strongly asymmetric spectra, with a marked peak at

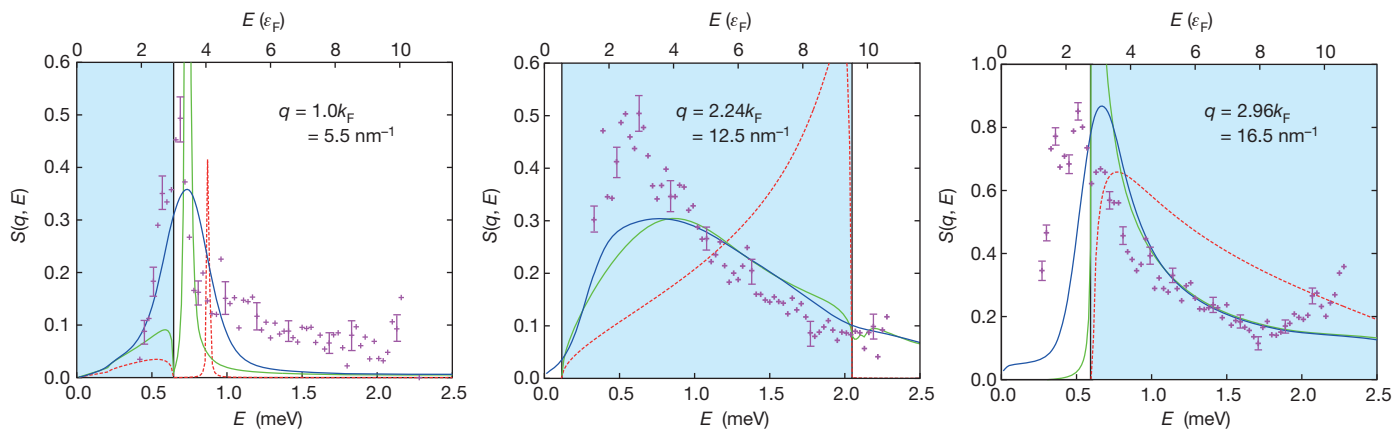


Figure 5 | Neutron spectra at selected wavevectors. The spectra correspond to cuts at wavevectors 5.5 , 12.5 and 16.5 nm^{-1} through the data shown in Fig. 3. Crosses, experimental data, with error bars calculated using the standard deviation of the neutron counts. (See Supplementary Information for details of the statistical sample.) Red dashed lines, results of RPA calculations (slightly

low energy. RPA calculations display qualitatively different asymmetric spectra, the peak being at high energy. This effect was previously seen in bulk ^3He measurements, impairing analysis of the spectra¹⁶. The dynamic many-body calculation⁴ that we apply here gives a much better description of the spectra.

The measured spectra are broadened, owing to the finite experimental resolution ($\Delta k = 1 \text{ nm}^{-1}$ and $\Delta E = 0.1 \text{ meV}$). The dynamic many-body calculation convolved with the experimental resolution provides calculated spectra whose widths agree with the measured ones, as seen in Fig. 5. Therefore, the energy width of the collective mode observed here outside the particle-hole band is certainly smaller than or equal to the energy resolution. This should not be confused with the extremely sharp phonon-roton mode of liquid ^4He in the superfluid state: liquid ^3He sustains, in addition to the collective density mode, incoherent particle-hole excitations, which open additional decay channels, as mentioned above.

A word is in order on our use of a non-interacting single-particle spectrum for determining the boundaries of the PHB. Incoherent particle-hole excitations are limited to a band $E(q - k) - E(k) < \hbar\omega < E(q + k) - E(k)$, where k is inside the Fermi sea. The single-particle spectrum $E(k)$ deviates from the free spectrum for two reasons. One is the corrugation of the substrate, leading to the so-called ‘band mass’, $m_{\text{band}}^* \approx 1.2m$ (ref. 26). Second, there is an interaction contribution. It exhibits a peak at the Fermi wavevector k_F , but decreases rapidly as a function of momentum in both three and two dimensions^{13–16}. As a consequence, the particle-hole band is modified only up to about the Fermi energy by interaction and band-structure effects.

In summary, we have observed the elementary excitations of two-dimensional liquid ^3He over a large range of energy and wavevector. Using the favourable conditions displayed by this system, we have demonstrated that a strongly interacting quantum many-body system sustains collective density excitations that are largely independent of the quantum statistics: the fermionic ^3He collective mode has the same physical origin as the phonon-roton curve of the bosonic ^4He . To gain a theoretical understanding of these phenomena, we developed a dynamical treatment of short-range correlations.

Generalizing Ruvalds’ proposal²⁷ of a superconducting pairing mechanism mediated by long-wavelength plasmons, our observation of a roton-like coherent mode characterized by a high density of states leads us to suggest a novel pairing mechanism, mediated by high-momentum density fluctuations. The consequences of the presence of plasmon collective modes at high wavevectors for the dynamics of electronic systems^{28–31}, including high- T_c superconductors, heavy fermions, metals and graphene, deserve exploration.

broadened to make delta-functions visible). Green lines, results from our dynamic many-body theory; blue lines, theoretical results convolved with the experimental resolution. The light blue area corresponds to the particle-hole band. The broadening of the results of the dynamic theory is due to multi-pair excitations.

Received 2 September 2011; accepted 1 February 2012.

- Landau, L. D. On the theory of superfluidity of helium II. *USSR J. Phys.* **11**, 91–92 (1947).
- Landau, L. D. The theory of a Fermi liquid. *Sov. Phys. JETP* **3**, 920–925 (1957).
- Pines, D. & Nozières, P. *The Theory of Quantum Liquids* (Benjamin, 1966).
- Böhm, H. M., Holler, R., Krotscheck, E. & Panholzer, M. Dynamic many-body theory: Dynamics of Strongly Correlated Fermi Fluids. *Phys. Rev. B* **82**, 224505 (2010).
- Thouless, D. J. *The Quantum Mechanics of Many-body Systems* 2nd edn (Academic, 1972).
- Pines, D. Elementary excitations in quantum liquids. *Phys. Today* **34**, 106–131 (1981).
- Feynman, R. P. & Cohen, M. Energy spectrum of the excitations in liquid helium. *Phys. Rev.* **102**, 1189–1204 (1956).
- Glyde, H. R. *Excitations in Liquid and Solid Helium* (Clarendon, 1994).
- Nozières, P. Is the roton in superfluid ^4He the ghost of a Bragg spot? *J. Low Temp. Phys.* **137**, 45–67 (2004).
- Jackson, H. W. Perturbative form of $S(k, \omega)$ for liquid ^4He : basic calculation and results. *Phys. Rev. A* **8**, 1529–1535 (1973).
- Clements, B. E., Krotscheck, E. & Tymczak, C. J. Multiphonon excitations in boson quantum films. *Phys. Rev. B* **53**, 12253–12275 (1996).
- Campbell, C. E. & Krotscheck, E. Dynamic many-body theory: pair fluctuations in bulk ^4He . *Phys. Rev. B* **80**, 174501 (2009).
- Friman, B. L. & Krotscheck, E. K. Zero sound, spin fluctuations, and effective mass in liquid ^3He . *Phys. Rev. Lett.* **49**, 1705–1708 (1982).
- Krotscheck, E., & Springer, J. Physical mechanisms for effective mass enhancement in ^3He . *J. Low Temp. Phys.* **132**, 281–295 (2003).
- Boronat, J., Casulleras, J. & Grau, V. Krotscheck, E. & Springer, J. Effective mass of two-dimensional ^3He . *Phys. Rev. Lett.* **91**, 085302 (2003).
- Glyde, H. R. *et al.* Effective mass, spin fluctuations and zero sound in liquid ^3He . *Phys. Rev. B* **61**, 1421–1432 (2000).
- Sköld, K., Pelizzari, C. A., Kleb, R. & Ostrowski, G. E. Neutron scattering study of elementary excitations in liquid helium-3. *Phys. Rev. Lett.* **37**, 842–845 (1976).
- Scherm, R. *et al.* Pressure dependence of elementary excitations in normal liquid helium-3. *Phys. Rev. Lett.* **59**, 217–220 (1987).
- Greywall, D. S. Heat capacity of multilayers of ^3He adsorbed on graphite at low millikelvin temperatures. *Phys. Rev. B* **41**, 1842–1862 (1990).
- Morhard, K.-D. *et al.* Two-dimensional Fermi liquid in the highly correlated regime: the second layer of ^3He adsorbed on graphite. *Phys. Rev. B* **53**, 2658–2661 (1996).
- Bäuerle, C., Bunkov, YuM, Chen, A. S., Fisher, S. N. & Godfrin, H. Ultra-low temperature magnetic properties of liquid ^3He films. *J. Low Temp. Phys.* **110**, 333–338 (1998).
- Casey, A., Patel, H., Nyéki, J., Cowan, B. P. & Saunders, J. Strongly correlated two dimensional fluid ^3He . *J. Low Temp. Phys.* **113**, 293–298 (1998).
- Neumann, M., Nyéki, J., Cowan, B. P. & Saunders, J. Bilayer ^3He : a simple two-dimensional heavy-fermion system with quantum criticality. *Science* **317**, 1356–1359 (2007).
- Godfrin, H. & Lauter, H. J. *Progress in Low Temperature Physics* Vol. XIV, Ch. 4, 213–320 (ed. Halperin, W. P.) (Elsevier Science, 1995).
- Lauter, H. J., Godfrin, H., Frank, V. L. P. & Leiderer, P. Ripplons in ^4He films observed by neutron scattering. *Phys. Rev. Lett.* **68**, 2484–2487 (1992).
- Reatto, L. Novel substrates for helium adsorption: graphane and graphene-fluoride. Lecture given at *Intl Conf. on Low Temp. Phys. (LT26)* (Beijing, 15 August 2011).
- Ruvalds, J. Are plasmons the key to superconducting oxides? *Nature* **328**, 299 (1987).
- Bostwick, A., Ohta, T., Seyller, T., Horn, K. & Rotenberg, E. Quasiparticle dynamics in graphene. *Nature Phys.* **3**, 36–40 (2007).
- Diaconescu, B. *et al.* Low-energy acoustic plasmons at metal surfaces. *Nature* **448**, 57–59 (2007).
- Uemura, Y. J. *et al.* Phase separation and suppression of critical dynamics at quantum phase transitions of MnSi and $(\text{Sr}_{1-x}\text{Ca}_x)\text{RuO}_3$. *Nature Phys.* **3**, 29–35 (2007).
- Kohsaka, Y. *et al.* How Cooper pairs vanish approaching the Mott insulator in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$. *Nature* **454**, 1072–1078 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to the Austrian–French programme Amadeus for providing initial support; to the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF grant P21264), the French Agence Nationale de la Recherche (project ANR-2010-INTB-403-01) and the EU FRP7 low-temperature infrastructure grant Microkelvin (project number 228464) for funding this research; and to the Institut Laue-Langevin for use of the facility.

Author Contributions H.G., M.M., H.-J.L., A.S. and M.P. performed the neutron experiments; H.M.B., E.K. and M.P. developed the theory. All authors analysed the results and contributed to the work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.G. (henri.godfrin@grenoble.cnrs.fr).

Experimental observation of electron–hole recollisions

B. Zaks¹, R. B. Liu² & M. S. Sherwin¹

An intense laser field can remove an electron from an atom or molecule and pull the electron into a large-amplitude oscillation in which it repeatedly collides with the charged core it left behind^{1–4}. Such recollisions result in the emission of very energetic photons by means of high-order-harmonic generation, which has been observed in atomic and molecular gases^{5–7} as well as in a bulk crystal⁸. An exciton is an atom-like excitation of a solid in which an electron that is excited from the valence band is bound by the Coulomb interaction to the hole it left behind^{9,10}. It has been predicted that recollisions between electrons and holes in excitons will result in a new phenomenon: high-order-sideband generation^{11,12}. In this process, excitons are created by a weak near-infrared laser of frequency f_{NIR} . An intense laser field at a much lower frequency, f_{THz} , then removes the electron from the exciton and causes it to recollide with the resulting hole. New emission is predicted to occur as sidebands of frequency $f_{\text{NIR}} + 2nf_{\text{THz}}$, where n is an integer that can be much greater than one. Here we report the observation of high-order-sideband generation in semiconductor quantum wells. Sidebands are observed up to eighteenth order ($+18f_{\text{THz}}$ or $n = 9$). The intensity of the high-order sidebands decays only weakly with increasing sideband order, confirming the non-perturbative nature of the effect. Sidebands are strongest for linearly polarized terahertz radiation and vanish when the terahertz radiation is circularly polarized. Beyond their fundamental scientific significance, our results suggest a new mechanism for the ultrafast modulation of light, which has potential applications in terabit-rate optical communications.

A recollision between an electron and an atomic core can be described by a three-step process¹³: first the electric field associated

with an intense laser ($\sim 10^{13} \text{ W cm}^{-2}$ at a wavelength of $\lambda = 1 \mu\text{m}$) ionizes an electron via tunneling¹⁴, then the ionized electron is quickly accelerated like a free particle away from and back to the atomic core by the optical field, and finally, on returning to the core, the electron recollides with the atomic core and emits radiation in the form of high harmonics of the laser frequency. Because tunneling predominantly occurs when the field intensity is at its maximum¹⁵, photons are emitted every half-period. The highest-order harmonic is determined by the maximum energy gained by the electron, $E_{\text{max}} \approx I_p + 3.2U_p$, where I_p is the ionization potential and $U_p = e^2 F^2 / 16 \pi^2 m_e f^2$ is the ponderomotive energy (F is the field strength, e is the elementary charge, m_e is the mass of the electron and f is the frequency of the optical field). By causing recollisions between optically excited electron–hole pairs, which, in gallium arsenide, have small effective masses ($m_e/15$) and millielectronvolt binding energies, it is predicted that high-order-sideband generation (HSG) can be achieved at much lower intensities ($\sim 10^5 \text{ W cm}^{-2}$) and with much lower-frequency fields^{11,12} ($f_{\text{THz}} \approx 0.5 \text{ THz}$, $\lambda_{\text{THz}} \approx 0.6 \text{ mm}$) than is necessary for high-order-harmonic generation (HHG) in atoms or molecules.

In this Letter, we describe recollisions between electrons and holes that result in HSG. We performed experiments on undoped semiconductor quantum well heterostructures because the confinement of electrons and holes enhances the Coulomb interaction and the formation of excitons relative to bulk material. Excitons are created in the quantum wells using a near-infrared (NIR) laser (Fig. 1a). The excitons are then driven by an intense terahertz field (generated by the free-electron lasers (FELs) at the Institute for Terahertz Science and Technology at the University of California, Santa Barbara), causing electron–hole recollisions and the generation of high-order sidebands (Fig. 1b).

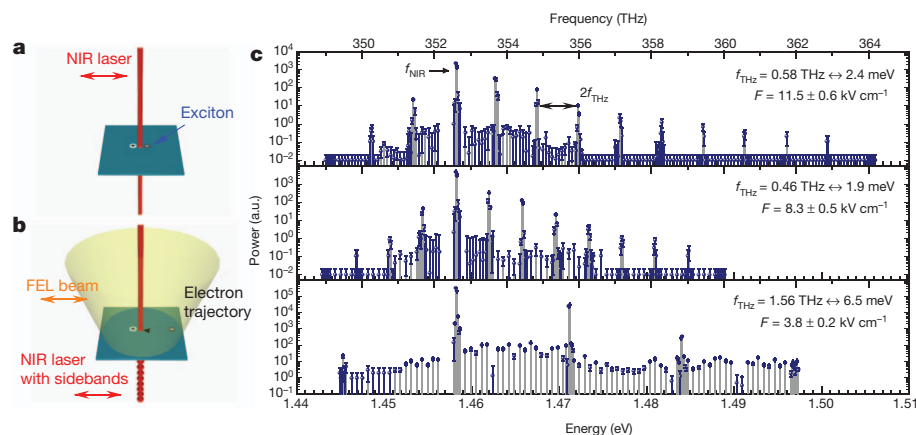


Figure 1 | Terahertz-sideband generation in a quantum well. **a**, Linearly polarized NIR light creates excitons when incident upon the quantum well sample. **b**, Intense, linearly polarized terahertz radiation can accelerate an electron away from and back to the exciton core, resulting in a recollision that emits light. HSG predominantly occurs during a short time near the peak of the electric field, such that sidebands are generated in short bursts separated by half the period of the terahertz radiation. Modulations of the NIR beam in the figure are representative of the intensity profile and are not shown to scale. **c**, Sideband spectra from a 15-nm-wide quantum well driven at various terahertz

frequencies. The signal from the NIR laser, of frequency f_{NIR} , is reduced by a factor of 500. Sidebands are observed at even multiples of the terahertz frequency, $f_{\text{sideband}} = f_{\text{NIR}} + 2nf_{\text{THz}}$, up to 18th order ($2n = 18$). At the two lower terahertz frequencies ($f_{\text{THz}} = 0.58$ and 0.46 THz), we observe a plateau in the sideband intensity for the high-order sidebands that indicates the non-perturbative nature of the generation process. At the highest terahertz frequency ($f_{\text{THz}} = 1.56 \text{ THz}$) and lower electric fields, the ponderomotive energy is smaller and many fewer sidebands are observed. a.u., arbitrary units. Error bars, s.e.m.

¹Department of Physics and the Institute for Terahertz Science and Technology, University of California at Santa Barbara, Santa Barbara, California 93106, USA. ²Department of Physics and Centre of Optical Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China.

HSG is observed at multiple driving frequencies. The observed sideband spectra of a 15-nm quantum well are shown for three driving frequencies in Fig. 1c. Owing to the inversion symmetry of the system, identical recollisions occur every half-period and sidebands of the NIR laser are observed only at even multiples of the FEL terahertz frequency: $f_{\text{sideband}} = f_{\text{NIR}} + 2nf_{\text{THz}}$, where $2n$ is the order of the sideband. At a frequency of 0.58 THz, sidebands of up to eighteenth order ($n = 9$) are observed. This phenomenon is robust: higher-order sidebands are observed in several quantum well samples and at temperatures up to 100 K (Supplementary Figs 1 and 2). As the NIR laser is detuned from the exciton resonance, the sideband intensity quickly decreases, demonstrating the excitonic nature of HSG (Supplementary Fig. 3). We note that the energy associated with the highest-order sideband, $18hf_{\text{THz}} = 43.2$ meV, is greater than the energy of the optical phonon in the material, $hf_{\text{LO}} \approx 36$ meV. The intensity of the sidebands decays weakly with order above $n = 3$. This behaviour is similar to the levelling off of intensity observed for high-order harmonics in HHG. The sideband plateau is also present when the FEL frequency is 0.45 THz, and sidebands of up to fourteenth order are observed. However, when the FEL frequency is increased to 1.56 THz, sidebands are generated only up to fourth order. This observation is consistent with the perturbative sideband generation seen in a quantum well exciton system at a similar frequency^{16,17}. Because the ponderomotive energy is lower at 1.56 THz and the 6.5-meV photon energy is comparable to the ~ 10 -meV binding energy, the semiclassical recollision model can no longer properly describe the sideband generation.

Excitonic nonlinear effects associated with intense terahertz fields^{18–23}, including perturbative sideband generation^{16,17,24–26}, have been observed previously. For sidebands generated perturbatively, the intensity of the sideband of order $2n$ scales as the driving intensity to the power of $2n$, I^{2n} (ref. 17). The plateau observed in the intensity of the high-order sidebands indicates that HSG is non-perturbative. This can be verified by investigating the dependence of individual sidebands on FEL intensity. Figure 2a shows the dependence of the $n = 1$ sideband ($f_{\text{sideband}} = f_{\text{NIR}} + 2f_{\text{THz}}$) on FEL power. At low power, we observe the expected perturbative behaviour: the sideband intensity scales quadratically with FEL power. As the FEL intensity is increased, the intensity dependence deviates from quadratic and the perturbative description breaks down. For the most intense FEL fields, the sideband intensity is constant and the sideband is saturated.

Although it is often difficult to detect nonlinear optical processes above second order, we observe sidebands of fourteenth order or higher at two FEL frequencies in our experiment. The FEL intensity dependence of the higher-order sidebands further verifies the non-perturbative nature of the generation process. Figure 2b shows the FEL intensity dependence of the fourth-, sixth- and eighth-order sidebands (blue dots) as well as a fit to the appropriate power law (black dashed lines). The data systematically deviate from the power-law scaling, implying that the sidebands are not being generated perturbatively.

Sidebands at frequencies less than f_{NIR} are observed at all FEL frequencies that we investigated. However, at most only three sidebands are observed and there is no plateau of higher-order sidebands. Unlike those of the sidebands at frequencies greater than f_{NIR} , the intensities of the sidebands below f_{NIR} decrease strongly with increasing order, implying that these sidebands are generated by means of a perturbative process (Supplementary Fig. 4). This is understandable because the terahertz field cannot impart negative kinetic energy to the optically excited carriers.

The dependence of sideband intensity on the ellipticity of the FEL driving fields supports the identification of electron–hole recollisions as the mechanism for HSG. Recollisions should be more likely to occur for the linear trajectories induced by linearly polarized driving fields than for the curved trajectories induced by elliptically polarized driving fields. Sideband generation is therefore expected to be least efficient for circularly polarized driving fields. We measured the sideband intensity as a function of the ellipticity of the FEL polarization. For a field with

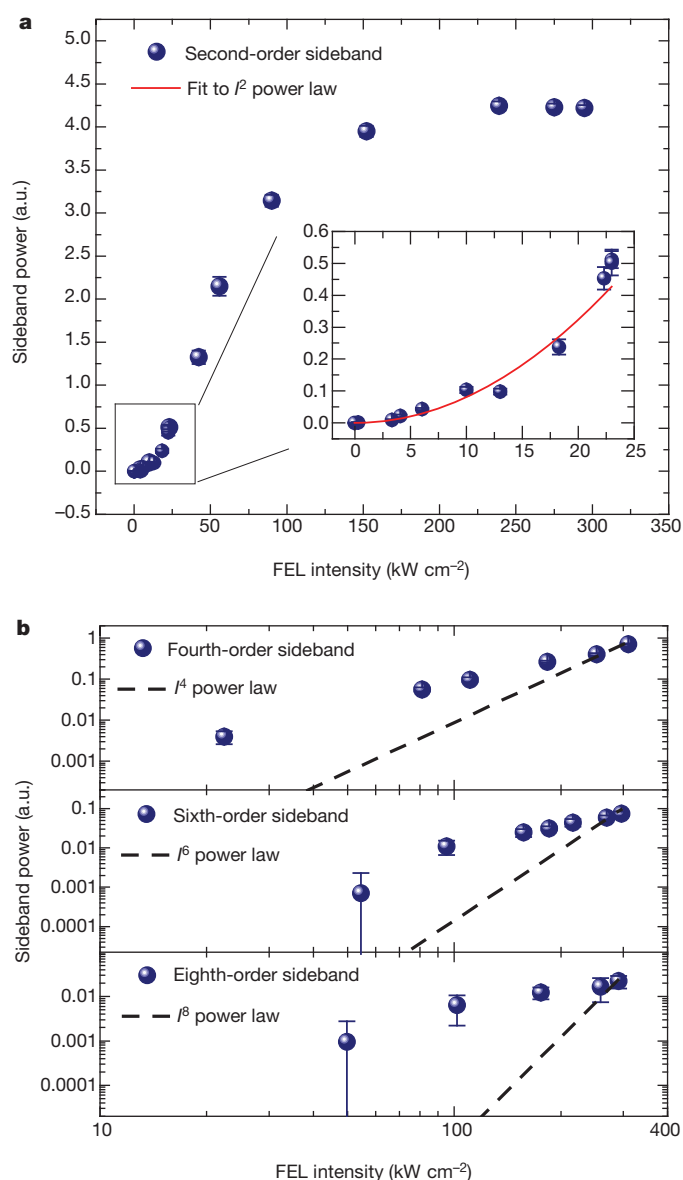


Figure 2 | Dependence of sideband strength on terahertz laser intensity. **a**, Dependence on FEL intensity of the second-order sideband ($n = 1$) produced in a 15-nm quantum well with $f_{\text{THz}} = 0.58$ THz. At low powers, the signal increases quadratically with FEL power (red line in inset), indicating a perturbative generation process. At high FEL intensities, the sideband saturates and the sideband signal is constant—behaviour that is non-perturbative. **b**, Dependence on FEL intensity of the fourth-, sixth- and eighth-order sidebands ($n = 2, 3$ and 4). A fit to the appropriate perturbative scaling law ($\propto I^{2n}$; black dashed lines) is plotted for each sideband. Our data systematically deviate from the I^{2n} fits, indicating that these sidebands are not generated perturbatively. Error bars, s.e.m.

the ellipticity given by the phase ϕ , the amplitudes of the fields in the x and y directions are given by $F_x = \cos(\phi/2)$ and $F_y = \sin(\phi/2)$, respectively. The dependence of the second-, fourth- and sixth-order sidebands ($n = 1, 2$ and 3) on the ellipticity of the FEL polarization is shown in Fig. 3 (blue dots). As predicted, the sideband intensity is at a maximum when the field is linearly polarized ($\phi = 0^\circ, 180^\circ, 360^\circ$) and a minimum when the field is circularly polarized ($\phi = 90^\circ, 270^\circ$). We also observed that the sideband intensity is slightly greater when the FEL polarization is parallel ($\phi = 0^\circ, 360^\circ$) to that of the NIR laser, rather than perpendicular ($\phi = 180^\circ$). Further investigation is necessary to understand the origin of this asymmetry.

The electrons responsible for the highest-order sidebands have the highest energy, travel the farthest from the exciton core and are the

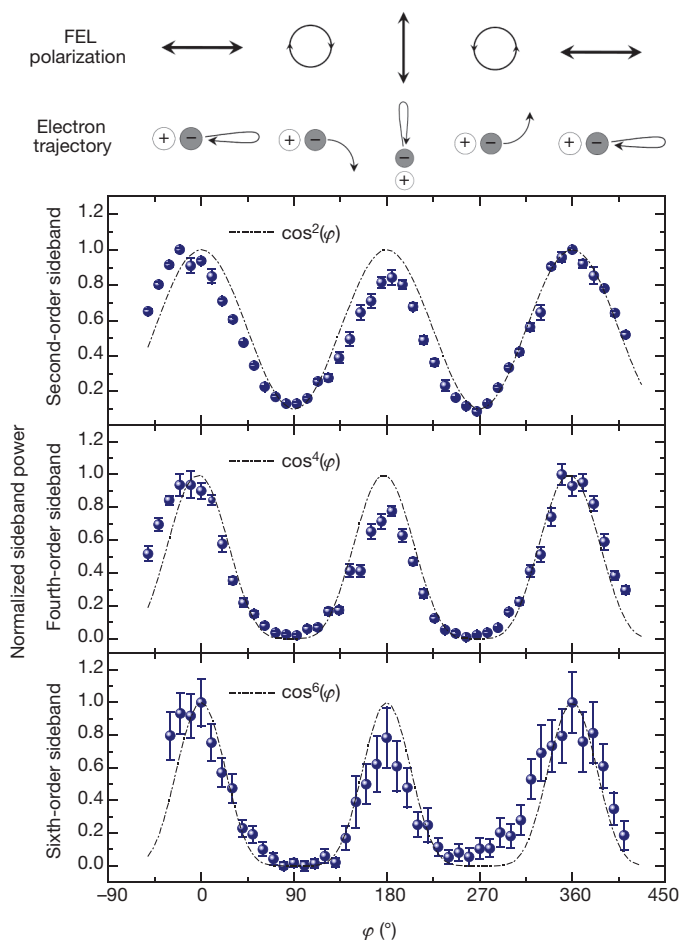


Figure 3 | Dependence of the sideband intensity on the ellipticity of the FEL polarization. The data are taken in a 15-nm quantum well with $f_{\text{THz}} = 0.58$ THz. The electric field amplitudes in the x and y directions are given by $F_x = \cos(\phi/2)$ and $F_y = \sin(\phi/2)$, respectively. With a circularly polarized field applied, the nonlinear trajectory of the electron has the lowest probability of recollision and sideband generation is minimized. By conservation of angular momentum, we expect the intensity of the n th sideband to scale as $\cos^{2n}(\phi)$. The predicted curves (black dash-dot line) are plotted with the experimental data (blue dots) for the second-, fourth-, and sixth-order sidebands ($n = 1, 2$ and 3). A sketch of the motion of the excited electron for a particular FEL polarization is shown at top. Error bars, s.e.m.

least likely to recollide when on a nonlinear path. Therefore, higher sidebands will have a stronger dependence on the ellipticity of the FEL polarization. A simple symmetry argument predicts that the intensity of the sideband of order $2n$ will scale as $\cos^{2n}(\phi)$ (Supplementary Equations). These curves are plotted in Fig. 3 (black dash-dot lines), and comparison with our data indicates that the dependence of the sideband intensity on the ellipticity of the driving fields is in good agreement with the predictions.

Like in HHG in atoms under intense illumination, the maximum energy released by the electron-hole recollision¹¹ is $E_{\text{max}} \approx \mathcal{A} + 3.2U_p$, where \mathcal{A} is the detuning from the band edge (here, the exciton binding energy) and the ponderomotive energy, $U_p = e^2 F^2 / 16 \pi^2 m^* f_{\text{THz}}^2$, is dependent on the FEL frequency, f_{THz} , the FEL field strength, F , and the effective electron-hole reduced mass, m^* . The nine sidebands observed at 0.58 THz are much fewer than the ~ 44 sidebands predicted by calculating $E_{\text{max}}/2\hbar f_{\text{THz}}$ assuming that $F = 11.5 \text{ kV cm}^{-1}$ and $m^* = 0.067m_e$. However, a number of effects that are neglected by the existing theory could limit the number of observable sidebands. Although the distance between excitons in our experiment is ~ 100 nm (Supplementary Discussion), the fields present suggest that the maximum excursion amplitude associated with the ponderomotive

motion is greater than 400 nm. Scattering of the highest-energy electrons may prevent the observation of the highest-order sidebands. Additionally, we expect that the maximum kinetic energy that can be reached by a ballistic electron in the indium gallium arsenide quantum wells should be close to the 36-meV threshold for optical phonon emission, which is slightly less than the 43.2 meV of excess energy carried by the highest-order sideband observed. Finally, the highest-energy sidebands we have observed are near the limits of detection for this experiment, so the highest-order sidebands may have escaped detection.

Observation of HSG demonstrates that the physics of recollisions extends beyond atoms and molecules. Though dielectric breakdown has made observation of HHG in solids difficult owing to the intense fields present during the short optical pulses, we have shown that recollisions in excitons can be studied at intensities well below the damage threshold of the semiconductor while using quasi-continuous wave sources. As is the case for recollisions in molecules^{27–29}, HSG may provide a unique tool for investigating the structure of optical excitations in condensed matter. Although the experiments described here used FELs as sources, the applied electric fields can actually be generated on-chip at frequencies greater than 300 GHz using existing transistor amplifiers³⁰. During the recollision process, the exciton is simultaneously coherent with both the NIR and the FEL fields, implying that recollisions controlled by the world's fastest transistors could provide an interesting platform for ultrafast modulation of light for optical communications.

METHODS SUMMARY

Excitons in $\text{In}_{0.06}\text{Ga}_{0.94}\text{As}/\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ quantum wells¹⁹ are created by NIR radiation (~ 350 THz) from a continuous-wave titanium-sapphire laser. The excitons are subjected to high-intensity radiation from FELs at frequencies between 0.4 and 1.6 THz. The beam from a FEL is focused onto the top of the quantum well sample such that the electric field is polarized in the plane of the quantum well. The NIR laser focus is overlapped with the FEL focus so that the beams are collinear when travelling through the sample, which is held at a temperature of 10 K. The sidebands are dispersed in a 0.85-m monochromator and detected with a photomultiplier tube.

Received 16 September 2011; accepted 13 January 2012.

- Krause, J. L., Schafer, K. J. & Kulander, K. C. High-order harmonic generation from atoms and ions in the high intensity regime. *Phys. Rev. Lett.* **68**, 3535–3538 (1992).
- Corkum, P. B. Plasma perspective on strong field multiphoton ionization. *Phys. Rev. Lett.* **71**, 1994–1997 (1993).
- Lewenstein, M., Balcou, P., Ivanov, M. Y., L'Huillier, A. & Corkum, P. B. Theory of high-harmonic generation by low-frequency laser fields. *Phys. Rev. A* **49**, 2117–2132 (1994).
- Corkum, P. B. Recollision physics. *Phys. Today* **64**, 36–41 (2011).
- Ferry, M. et al. Multiple-harmonic conversion of 1064 nm radiation in rare gases. *J. Phys. B* **21**, L31–L35 (1988).
- Liang, Y., Augst, S., Chin, S. L., Beaudoin, Y. & Chaker, M. High harmonic generation in atomic and diatomic molecular gases using intense picosecond laser pulses—a comparison. *J. Phys. B* **27**, 5119–5130 (1994).
- Krausz, F. & Ivanov, M. Attosecond physics. *Rev. Mod. Phys.* **81**, 163–234 (2009).
- Ghimire, S. et al. Observation of high-order harmonic generation in a bulk crystal. *Nature Phys.* **7**, 138–141 (2011).
- Wannier, G. H. The structure of electronic excitation levels in insulating crystals. *Phys. Rev.* **52**, 191–197 (1937).
- Dingle, R., Wiegmann, W. & Henry, C. H. Quantum states of confined carriers in very thin $\text{Al}_x\text{Ga}_{1-x}\text{As-GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ heterostructures. *Phys. Rev. Lett.* **33**, 827–830 (1974).
- Liu, R.-B. & Zhu, B.-F. High-order THz-sideband generation in semiconductors. *AIP Conf. Proc.* **893**, 1455–1456 (2007).
- Yan, J.-Y. Theory of excitonic high-order sideband generation in semiconductors under a strong terahertz field. *Phys. Rev. B* **78**, 075204 (2008).
- Corkum, P. B. & Krausz, F. Attosecond science. *Nature Phys.* **3**, 381–387 (2007).
- Keldysh, L. V. Ionization in the field of a strong electromagnetic wave. *Sov. Phys. JETP* **20**, 1307–1314 (1965).
- Uiberacker, M. et al. Attosecond real-time observation of electron tunnelling in atoms. *Nature* **446**, 627–632 (2007).
- Wagner, M. et al. Resonant enhancement of second order sideband generation for intraexcitonic transitions in GaAs/AlGaAs multiple quantum wells. *Appl. Phys. Lett.* **94**, 241105 (2009).

17. Wagner, M. *et al.* Terahertz nonlinear optics using intra-excitonic quantum well transitions: sideband generation and AC Stark splitting. *Phys. Status Solidi* **248**, 859–862 (2011).
18. Nordstrom, K. B. *et al.* Excitonic dynamical Franz-Keldysh effect. *Phys. Rev. Lett.* **81**, 457–460 (1998).
19. Carter, S. G. *et al.* Quantum coherence in an optical modulator. *Science* **310**, 651–653 (2005).
20. Danielson, J. R. *et al.* Interaction of strong single-cycle terahertz pulses with semiconductor quantum wells. *Phys. Rev. Lett.* **99**, 237401 (2007).
21. Leinß, S. *et al.* Terahertz coherent control of optically dark paraexcitons in Cu₂O. *Phys. Rev. Lett.* **101**, 246401 (2008).
22. Wagner, M. *et al.* Observation of the intraexciton Autler-Townes effect in GaAs/AlGaAs semiconductor quantum wells. *Phys. Rev. Lett.* **105**, 167401 (2010).
23. Zaks, B. *et al.* THz-driven quantum wells: Coulomb interactions and Stark shifts in the ultrastrong coupling regime. *N. J. Phys.* **13**, 083009 (2011).
24. Nordstrom, K. B. *et al.* Observation of dynamical Franz-Keldysh effect. *Phys. Status Solidi* **204**, 52–54 (1997).
25. Kono, J. *et al.* Resonant terahertz optical sideband generation from confined magnetoexcitons. *Phys. Rev. Lett.* **79**, 1758–1761 (1997).
26. Černe, J. *et al.* Near-infrared sideband generation induced by intense far-infrared radiation in GaAs quantum wells. *Appl. Phys. Lett.* **70**, 3543–3545 (1997).
27. Itatani, J. *et al.* Tomographic imaging of molecular orbitals. *Nature* **432**, 867–871 (2004).
28. Haessler, S. *et al.* Attosecond imaging of molecular electronic wavepackets. *Nature Phys.* **6**, 200–206 (2010).
29. Vozzi, C. *et al.* Generalized molecular orbital tomography. *Nature Phys.* **7**, 822–826 (2011).
30. Eisele, H. State of the art and future of electronic sources at terahertz frequencies. *Electron. Lett.* **46**, S8–S11 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Samples are from the wafers grown at UCSB by C. Wang in the group of L. C. Coldren (15-nm quantum well) and by T. Truong in the group of P. M. Petroff (18- and 22-nm quantum wells; see Supplementary Information) for experiments described in refs 19 and 23, respectively. We would like to thank D. Enyeart for his help in running the UCSB FELs and N. H. Kwong for running simulations helpful to our understanding of these effects. This work was supported by NSF grant DMR-1006603 and Hong Kong RGC/GRF 401011.

Author Contributions B.Z. was responsible for conducting experiments and analysing the data presented, and for writing the paper. R.B.L. introduced M.S.S. to the predictions of HSG, and provided theoretical support. M.S.S. designed and supervised the study and edited the paper. All authors discussed the results and commented on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.S.S. (sherwin@physics.ucsb.edu).

Deciphering a neuronal circuit that mediates appetite

Qi Wu^{1†}, Michael S. Clark² & Richard D. Palmiter¹

Hypothalamic neurons that co-express agouti-related protein (AgRP), neuropeptide Y and γ -aminobutyric acid (GABA) are known to promote feeding and weight gain by integration of various nutritional, hormonal, and neuronal signals^{1,2}. Ablation of these neurons in mice leads to cessation of feeding that is accompanied by activation of *Fos* in most regions where they project^{3–6}. Previous experiments have indicated that the ensuing starvation is due to aberrant activation of the parabrachial nucleus (PBN) and it could be prevented by facilitating GABA_A receptor signalling in the PBN within a critical adaptation period⁵. We speculated that loss of GABA signalling from AgRP-expressing neurons (AgRP neurons) within the PBN results in unopposed excitation of the PBN, which in turn inhibits feeding. However, the source of the excitatory inputs to the PBN was unknown. Here we show that glutamatergic neurons in the nucleus tractus solitarius (NTS) and caudal serotonergic neurons control the excitability of PBN neurons and inhibit feeding. Blockade of serotonin (5-HT₃) receptor signalling in the NTS by either the chronic administration of ondansetron or the genetic inactivation of *Tph2* in caudal serotonergic neurons that project to the NTS protects against starvation when AgRP neurons are ablated. Likewise, genetic inactivation of glutamatergic signalling by the NTS onto *N*-methyl D-aspartate-type glutamate receptors in the PBN prevents starvation. We also show that suppressing glutamatergic output of the PBN reinstates normal appetite after AgRP neuron ablation, whereas it promotes weight gain without AgRP neuron ablation. Thus we identify the PBN as a hub that integrates signals from several brain regions to bidirectionally modulate feeding and body weight.

Administration of diphtheria toxin (DT) to *Agrp^{DTR}* mice, which express the human DT receptor selectively in AgRP neurons, ablates nearly all AgRP neurons in the arcuate nucleus of the hypothalamus; during the next 6 days the mice gradually cease eating, lose body weight, and die without intervention⁴. Chronic infusion of bretazenil, a partial agonist of GABA_A receptor, into the PBN during the DT treatment prevents starvation and allows an adaptive process to take place such that the mice eat and maintain their body weight⁵. Not only does ablation of AgRP neurons inhibit the initiation of meals, it also decreases the amount of liquid food that will be swallowed when it is delivered directly into the mouth⁷. Because the PBN responds to visceral malaise, such as food poisoning and LiCl treatment⁸, and also processes gustatory signals in paradigms such as the conditional taste aversion or preference^{9,10}, we predicted that ablation of AgRP neurons results in unopposed activation of the PBN, which may mimic a nausea signal and thereby inhibit feeding. To test this hypothesis, we infused ondansetron, an anti-nausea drug that antagonizes 5-HT₃ receptors¹¹, subcutaneously or directly into the fourth ventricle, starting 3 days before injection of *Agrp^{DTR}* mice with DT. Despite the fact that the drug is administered orally to people, only central delivery of ondansetron prevented fatal weight loss and allowed the mice to recover (Fig. 1a and Supplementary Fig. 1a). Consumption of low-fat chow pellets by ondansetron-treated mice fell and they lost roughly 10% of their body weight during the first week after DT treatment; however, they then gradually ate more and

regained body weight by 3 weeks after DT treatment (Fig. 1a and Supplementary Fig. 1a). The 5-HT₃ receptor is an excitatory ion channel that is expressed widely in the brain, especially in the cortex and dorsal brainstem¹². To examine more precisely where ondansetron acts to prevent starvation after AgRP neuron ablation, the drug was delivered bilaterally to either the PBN or the NTS (see Supplementary Fig. 2 for cannula placement). Delivery of ondansetron to the PBN did not rescue the starvation phenotype of DT-treated mice, whereas delivery to the NTS prevented starvation (Fig. 1b and Supplementary Fig. 1b). The results suggest that serotonin provides some of the excitatory drive that indirectly results in hyperactivity of the PBN after loss of inhibitory input from AgRP neurons. Neurons in the NTS are known to send excitatory, glutamatergic inputs to the PBN^{13,14}. We therefore predicted that serotonin action on 5-HT₃ receptors in the NTS promotes hyperexcitation of the PBN, which can be measured as local *Fos* gene activation⁶. Consistent with this hypothesis, *Fos* induction in the PBN was significantly ameliorated by the administration of ondansetron in the NTS (Supplementary Fig. 3). We conclude that inhibition of 5-HT₃-mediated excitatory currents in the vicinity of the NTS prevents starvation after the ablation of AgRP neurons and promotes an adaptation that allows feeding to be maintained in the absence of AgRP neurons.

Tryptophan hydroxylase 2 (*Tph2*) catalyses the first and rate-limiting step in serotonin biosynthesis in the central nervous system¹⁵. To examine the role of serotonin more directly, conditional *Tph2^{lox/lox}* mice carrying the *Agrp^{DTR/+}* allele were generated and then injected with *CAV2-Cre*, a virus that is retrogradely transported from the site of injection to the cell bodies where it can inactivate the *Tph2* gene only in those neurons that project to the injection site. *CAV2-Cre* was injected bilaterally into the NTS of *Tph2^{lox/lox}*, *Agrp^{DTR/+}* mice, and 8 days later they were treated with DT to ablate the AgRP neurons. This viral treatment prevented the starvation that normally occurs after ablation of AgRP neurons. Feeding and body weight decreased slightly after DT treatment of the virally transduced mice, but some of the mice restored normal food intake and regained body weight (Fig. 1c and Supplementary Fig. 1c). Various raphe nuclei from the virally rescued *Tph2^{lox/lox}* mice were examined for serotonergic neurons that lacked *Tph2* but retained L-aromatic amino acid decarboxylase (AADC), another marker of serotonergic cells. Many serotonergic cell bodies in the raphe obscurus (ROb) and raphe magnus (RMg) were found that significantly lacked *Tph2* staining (Fig. 2a–l). Quantification of the results revealed that viral treatment decreased *Tph2* signalling in ROb and RMg serotonergic neurons by 60–80% (Fig. 2m), whereas serotonergic neurons in the dorsal raphe (DR) were unaffected (Fig. 2m, and data not shown). Viral injection into the NTS of *Tph2^{lox/lox}* mice decreased serotonin levels in the NTS by about 75% (wild-type, 11.65 ± 0.92 ng per mg of protein; virus-injected, 2.68 ± 0.57 ng per mg of protein). Our observations are consistent with known projections of the caudal raphe neurons to brainstem structures and projections of DR neurons to forebrain regions¹⁶.

Because many of the neurons in the NTS are known to send glutamatergic projections to the PBN^{13,14}, we predicted that serotonergic

¹Howard Hughes Medical Institute and Department of Biochemistry, University of Washington School of Medicine, Seattle, Washington 98195, USA. ²Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA. [†]Present address: Eagles Diabetes Research Center and Department of Pharmacology, Carver College of Medicine, University of Iowa, Iowa City, Iowa 52242, USA.

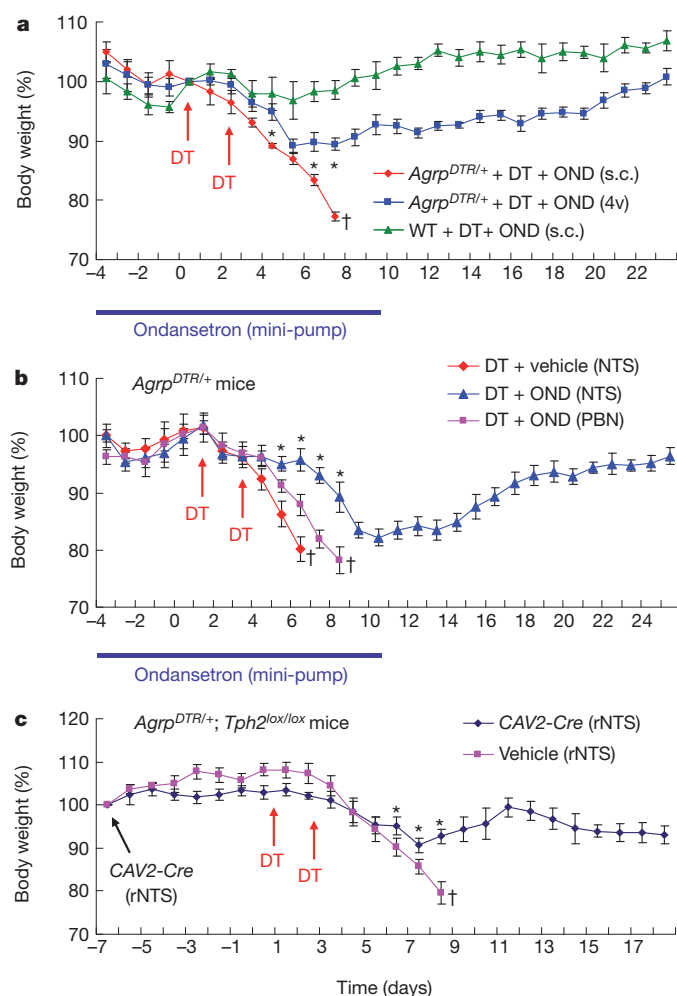


Figure 1 | Chronic administration of ondansetron into the NTS, or genetic inactivation of serotonergic input to the NTS prevents starvation in AgRP neuron-ablated mice. **a**, Body weight of *AgRP^{DTR/+}* mice and wild-type (WT) mice after either subcutaneous (s.c.; $1 \text{ mg kg}^{-1} \text{ d}^{-1}$, $n = 6$) or fourth-ventricle (4v; $0.1 \text{ mg kg}^{-1} \text{ d}^{-1}$, $n = 12$) infusion of ondansetron, a 5-HT₃ receptor antagonist, through osmotic minipumps. DT was injected intramuscularly twice as indicated by arrows. Minipumps were removed on day 11 after the pump content was depleted. Eight of 12 mice survived the treatment of DT and infusion of ondansetron into 4v (blue line). Asterisk, $P < 0.01$ between *AgRP^{DTR/+}* mice treated with ondansetron s.c. ($n = 6$ non-survivors) and 4v ($n = 8$ survivors). **b**, Body weight of DT-treated *AgRP^{DTR/+}* mice after chronic infusion of ondansetron or vehicle into either the NTS ($n = 14$) or the PBN ($n = 8$). DT was twice injected intramuscularly as indicated by arrows. Six of 14 mice survived the treatment of DT and infusion of ondansetron into the NTS (blue line). Asterisk, $P < 0.01$ between *AgRP^{DTR/+}* mice infused by ondansetron into the NTS ($n = 6$ survivors) and the PBN ($n = 8$ non-survivors). **c**, Body weight of DT-treated, *AgRP^{DTR/+}; Tph2^{lox/lox}* mice after bilateral injection of CAV2-Cre virus or vehicle into the rostral part of the NTS (rNTS). Seven of 12 mice survived the treatment with DT and viral infection (blue line). Asterisk, $P < 0.05$ between mice injected with CAV2-Cre ($n = 7$ survivors) and vehicle into the NTS ($n = 8$ non-survivors). Daggers indicate mice that were removed from the experiment when they either lost 20% of their body weight or seemed moribund. Results are shown as means \pm s.e.m.

activation of the NTS might lead to glutamatergic activation of the PBN, which could be responsible for starvation after AgRP neuron ablation. To test this hypothesis, we inactivated the vesicular glutamate transporter 2 (Vglut2, encoded by the *Slc17a6* gene) within the NTS by injecting AAV1-CreGFP virus bilaterally into the NTS of *Slc17a6^{lox/lox}; AgRP^{DTR/+}* mice 8 days before the initiation of DT treatment (see Supplementary Fig. 4 for the placement of virus injections). Mice that were injected bilaterally and treated with DT maintained body weight and feeding, whereas mice that were unilaterally injected starved, as did mice with vehicle

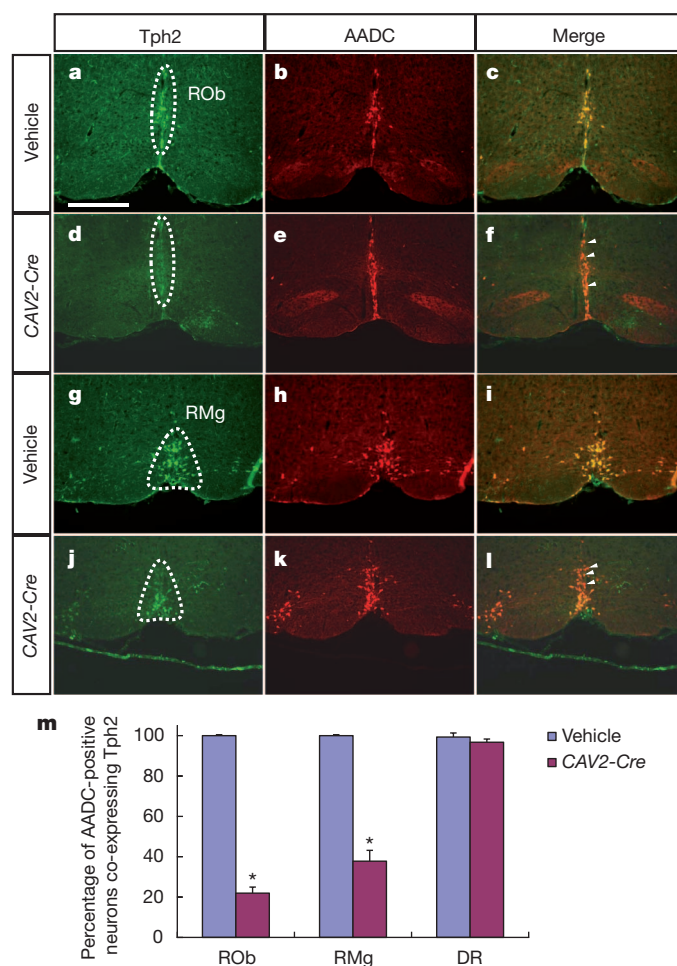


Figure 2 | Serotonergic projections from the ROB and RMg to the NTS mediate starvation after ablation of AgRP neurons. **a–f**, Representative immunohistochemistry pictures of Tph2 and AADC, markers of serotonergic neurons, in *AgRP^{DTR/+}; Tph2^{lox/lox}* mice after bilateral injection of either vehicle (**a–c**) or a retrograding CAV2-Cre virus (**d–f**) into the rostral NTS. Arrowheads indicate the expression of Tph2 was gone after viral infection. **g–l**, Representative immunostaining pictures of Tph2 and AADC at the RMg (B3 group) from the mice described above. Arrowheads indicate the serotonin neurons within the RMg in which the expression of Tph2 was gone after viral infection. **m**, Quantified immunohistochemistry results of AADC-expressing neurons that co-localized with Tph2-expressing neurons in the ROB, RMg and DR of the mice described in **a–f** (and data not shown). Asterisk, $P < 0.01$, analysis of variance (ANOVA); $n = 6$ mice per group. Results are shown as means and s.e.m. Scale bar in **a** (for **a–l**), 400 μm .

injection (Fig. 3a and Supplementary Fig. 5). To further establish that glutamatergic activation of the PBN inhibits feeding in this model, we used a viral/genetic approach to reduce N-methyl-D-aspartate (NMDA) receptors in the PBN and thereby dampen the excitability by glutamate. *Grin1^{lox/lox}; AgRP^{DTR/+}* mice, which carry two conditional alleles of the gene encoding the essential NR1 subunit of NMDA receptors and the *AgRP^{DTR}* allele, were injected bilaterally with AAV1-CreGFP in the PBN 9 days before ablation of the AgRP neurons by DT (see Supplementary Fig. 6 for the placement of virus injections). For the bilaterally injected mice, body weight declined slightly during the first 8 days after DT injection, but then recovered along with a rebound of food intake, whereas the vehicle-injected and unilaterally injected mice stopped eating and did not recover (Fig. 3b and Supplementary Fig. 7). These experiments show that either decreasing glutamatergic signalling by neurons within the NTS or decreasing the number of NMDA receptors in the PBN protects against the starvation caused by the ablation of AgRP neurons. Most neurons within the PBN are glutamatergic. We therefore predicted that

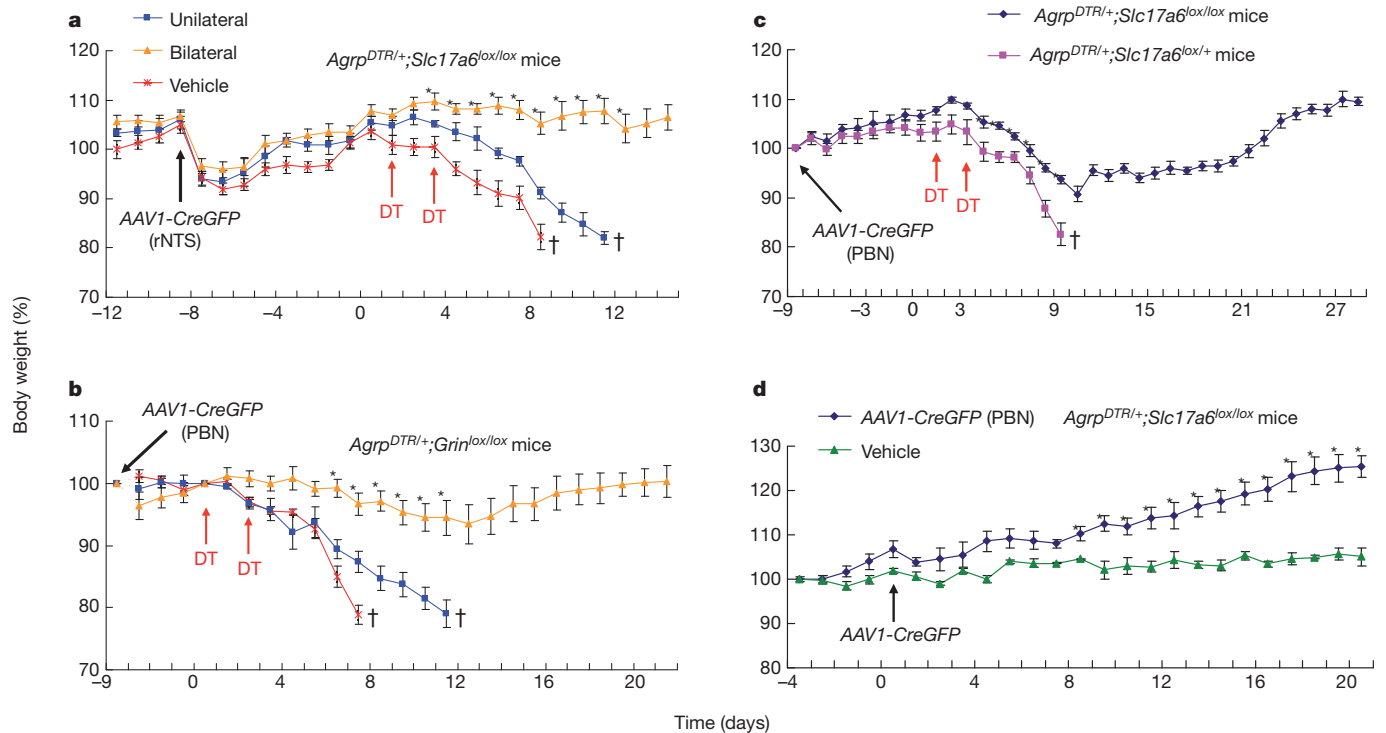


Figure 3 | Viral-mediated disruption of glutamatergic circuitry between the NTS and PBN, or glutamatergic output of the PBN, rescues feeding after ablation of AgRP neurons. **a**, Percentage of initial body weight of DT-treated *AgRP^{DTR/+}; Slc17a6^{lox/lox}* mice after either bilateral ($n = 19$) or unilateral ($n = 9$) injection of AAV1-CreGFP virus or vehicle into the rostral NTS (rNTS; $n = 9$). AAV1-CreGFP virus decreases glutamatergic signalling from the rNTS to downstream targets, including the PBN. Out of 19 mice, 14 survived DT treatment and viral injections into the NTS (orange line). Asterisk, $P < 0.01$ between bilateral virus injection ($n = 14$ survivors) and unilateral virus injection ($n = 9$ non-survivors) or vehicle injection ($n = 9$ non-survivors). **b**, Percentage of initial body weight of DT-treated *AgRP^{DTR/+}; Grin1^{lox/lox}* mice after either bilateral injection ($n = 18$) or unilateral injection of AAV1-CreGFP virus ($n = 8$) or vehicle ($n = 8$) into the lateral PBN. AAV1-CreGFP virus attenuates NMDA receptor signalling in the PBN, which receives dense glutamatergic projections from the NTS. Out of 18 mice, 11 survived DT treatment and viral injections into the PBN (orange line). Asterisk, $P < 0.01$

suppression of glutamatergic signalling by the PBN should also protect against starvation when AgRP neurons are ablated. We used the *Slc17a6^{lox/lox}; AgRP^{DTR/+}* mice again for this experiment and injected AAV1-CreGFP into the lateral PBN 9 days before DT treatment (Supplementary Fig. 5). The virally injected mice lost body weight during the first week after DT treatment but then gradually recovered, whereas the controls did not recover (Fig. 3c and Supplementary Fig. 8). In another cohort of mice, AAV1-CreGFP was injected into the PBN of *Slc17a6^{lox/lox}; AgRP^{DTR/+}* mice, but they were not treated with DT. Those mice gained about 20% in body weight which was accompanied by about 16% increase in food intake over the next 3 weeks (Fig. 3d and Supplementary Fig. 9). Our results indicate that enhanced glutamatergic signalling by the PBN inhibits feeding and promotes weight loss, whereas lowering the glutamatergic output of the PBN promotes weight gain through an increase in feeding, possibly combined with a decrease in energy expenditure.

Our studies reveal six manipulations that allow mice to survive after the ablation of AgRP neurons: enhancement of GABA_A receptor signalling in the PBN with bretazenil⁵; suppression of 5-HT₃ receptor signalling in the NTS with ondansetron; disablement of serotonergic input to the NTS through viral-mediated removal of Tph2; reduction of glutamatergic activation of the PBN by decreasing the number of NMDA receptors; or reduction of efferent glutamatergic

between bilateral virus injection ($n = 11$ survivors) and unilateral virus injection ($n = 8$ non-survivors) or vehicle injection ($n = 8$ non-survivors). **c**, Percentage of initial body weight of DT-treated *AgRP^{DTR/+}; Slc17a6^{lox/lox}* mice ($n = 21$) or DT-treated *AgRP^{DTR/+}; Slc17a6^{lox/+}* control mice ($n = 10$) after bilateral injection of AAV1-CreGFP virus into the lateral PBN. AAV1-CreGFP virus abolishes glutamatergic signalling from the PBN to various forebrain targets. Out of 21 mice, 12 survived DT treatment and viral injections into the PBN (blue lines). Asterisk, $P < 0.01$, ANOVA between 12 survivors of Vglut2-deficient group and 10 non-survivor control group. **d**, Percentage of initial body weight of *AgRP^{DTR/+}; Slc17a6^{lox/lox}* mice after bilateral injection of AAV1-CreGFP virus ($n = 14$) or vehicle ($n = 10$) into the PBN. Asterisk, $P < 0.01$; ANOVA between mice with precise viral injection ($n = 7$) and vehicle injection ($n = 10$). Results are shown as means \pm s.e.m. Daggers indicate mice that were removed from experiment when they either lost more than 20% of body weight or seemed moribund.

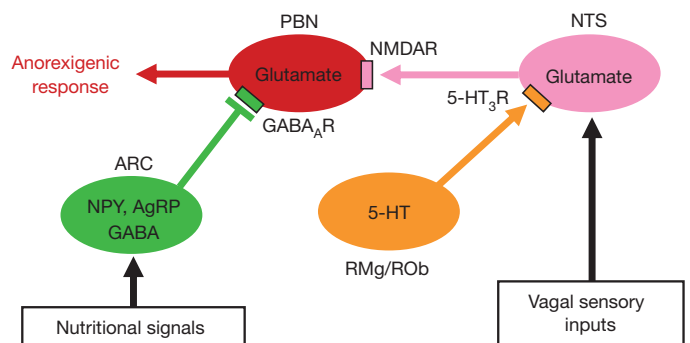


Figure 4 | Diagram illustrating circuitry that mediates loss of appetite after acute ablation of hypothalamic AgRP neurons. AgRP neurons co-expressing AgRP, neuropeptide Y (NPY) and GABA send inhibitory projections to the PBN. Serotonergic neurons residing in the RMg and ROb inhibit feeding through excitation of postsynaptic neurons in rostral NTS that express 5-HT₃ receptors (5-HT₃R). This subpopulation of NTS neurons, by integration of visceral and gustatory inputs, sends excitatory glutamate signalling to the lateral PBN neurons that express NMDA receptors (NMDAR). Nutritional signalling from the hypothalamus and sensory signals may interact within the PBN to promote appropriate feeding responses. ARC, arcuate nucleus; GABA_A receptor.

signals from the PBN. These observations support the circuit depicted in Fig. 4. We suggest that a subpopulation of neurons in the PBN integrates visceral and gustatory information from the NTS with energy-balance signals emanating from AgRP neurons. The NTS responds to vagal inputs as well as gut-derived hormones, while AgRP neurons detect nutrient levels and respond to hormones such as insulin, leptin and ghrelin^{17–19}. Consequently, the appetitive response can be modulated by food palatability and visceral condition in a manner dictated by current energy balance.

Some studies have suggested that serotonin exerts its anorectic effects by differential actions on 5-HT_{1b} and 5-HT_{2c} receptors in the hypothalamus to stimulate melanocortin signalling^{20–22}, whereas a recent study indicated that serotonergic neurons in ventral raphe nuclei respond to food restriction by an elevated Fos signal²³. We show here that serotonin from the ROb and RMg acts on 5-HT₃ receptors in the NTS to inhibit feeding after the ablation of AgRP neurons; thus, some of the anorectic effects of serotonin reuptake inhibitors, such as fenfluramine, may also be mediated in the brainstem. Classical mapping studies reveal projections from the PBN to the amygdala, thalamus, hypothalamus and other brain regions²⁴. Further characterization and manipulation of PBN circuits that control feeding will be greatly facilitated by the identification of genes that are expressed exclusively by the relevant subpopulations of PBN neurons. Our experiments help define an important neural pathway within which some unique therapeutic targets have been characterized that could be valuable for the development of new treatments of various eating disorders, including nausea and anorexia nervosa^{25,26}.

METHODS SUMMARY

AgRP^{DTR/+} mice, *Grin1*^{lox/lox} mice and *Slc17a6*^{lox/lox} mice on a C57Bl/6 background were generated and identified genetically by PCR of tail DNA as described previously^{4,27,28}. A conditional *Tph2* targeting construct was prepared by flanking the first exon with *loxP* sites; the generation of *Tph2*^{lox/lox} mice on a mixed C57Bl/6 x 129/Sv background was based on standard protocols. All experiments were performed with male mice about 8 weeks old. To ablate AgRP neurons, mice carrying the *AgRP*^{DTR} allele were injected twice intramuscularly with DT (50 µg kg⁻¹; List Biological Laboratories, Campbell); the injections were 2 days apart⁴. The extent of ablation (more than 95%) was determined by immunohistochemistry^{4–7}. Production of adeno-associated virus 1 (AAV1-CreGFP) and canine adenovirus 2 (CAV2-Cre) followed the protocols described previously^{29,30}. For injection of virus into the NTS or PBN, mice were anaesthetized and virus was injected bilaterally or unilaterally through a 5-µl Hamilton syringe. For drug treatments, Alzet 14-day minipumps (model 1002) loaded with 100 µl of ondansetron (6 mg ml⁻¹; Sigma-Aldrich) were implanted subcutaneously on the back of anaesthetized mice 4 days before DT treatment. Alternatively, cannulas (28 gauge; Plastics One) were placed into fourth ventricles and connected by tubing with the subcutaneous minipumps that were loaded with ondansetron (0.6 mg ml⁻¹). For feeding assays, mice were transferred to BioDAQ Food and Water Intake Monitor (Research Diets) supplied with water and low-fat chow diet (D12450B). The mice were allowed to acclimatize for 3 days before the initiation of each experiment and data collection. Body weight and total food intake were recorded every 24 h. Feeding and drinking activity was recorded in accordance with the manufacturer's suggested protocol.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 July 2011; accepted 24 January 2012.

Published online 14 March 2012.

- Wu, Q. & Palmiter, R. D. GABAergic signaling by AgRP neurons prevents anorexia via a melanocortin-independent mechanism. *Eur. J. Pharmacol.* **660**, 21–27 (2011).
- Morton, G. J., Cummings, D. E., Baskin, D. G., Barsh, G. S. & Schwartz, M. W. Central nervous system control of food intake and body weight. *Nature* **443**, 289–295 (2006).
- Gropp, E. *et al.* Agouti-related peptide-expressing neurons are mandatory for feeding. *Nature Neurosci.* **8**, 1289–1291 (2005).
- Luquet, S., Perez, F. A., Hnasko, T. S. & Palmiter, R. D. NPY/AgRP neurons are essential for feeding in adult mice but can be ablated in neonates. *Science* **310**, 683–685 (2005).

- Wu, Q., Boyle, M. P. & Palmiter, R. D. Loss of GABAergic signaling by AgRP neurons to the parabrachial nucleus leads to starvation. *Cell* **137**, 1225–1234 (2009).
- Wu, Q., Howell, M. P. & Palmiter, R. D. Ablation of neurons expressing agouti-related protein activates Fos and gliosis in postsynaptic target regions. *J. Neurosci.* **28**, 9218–9226 (2008).
- Wu, Q., Howell, M. P., Cowley, M. A. & Palmiter, R. D. Starvation after AgRP neuron ablation is independent of melanocortin signaling. *Proc. Natl Acad. Sci. USA* **105**, 2687–2692 (2008).
- Swank, M. W. & Bernstein, I. L. c-Fos induction in response to a conditioned stimulus after single trial taste aversion learning. *Brain Res.* **636**, 202–208 (1994).
- Yamamoto, T. Neural substrates for the processing of cognitive and affective aspects of taste in the brain. *Arch. Histol. Cytol.* **69**, 243–255 (2006).
- Berridge, K. C. & Pecina, S. Benzodiazepines, appetite, and taste palatability. *Neurosci. Biobehav. Rev.* **19**, 121–131 (1995).
- Gershon, M. D. & Tack, J. The serotonin signaling system: from basic understanding to drug development for functional GI disorders. *Gastroenterology* **132**, 397–414 (2007).
- Barnes, N. M., Hales, T. G., Lummis, S. C. & Peters, J. A. The 5-HT₃ receptor—the relationship between structure and function. *Neuropharmacology* **56**, 273–284 (2009).
- Herbert, H., Moga, M. M. & Saper, C. B. Connections of the parabrachial nucleus with the nucleus of the solitary tract and the medullary reticular formation in the rat. *J. Comp. Neurol.* **293**, 540–580 (1990).
- Jhamandas, J. H. & Harris, K. H. Excitatory amino acids may mediate nucleus tractus solitarius input to rat parabrachial neurons. *Am. J. Physiol.* **263**, R324–R330 (1992).
- Walther, D. J. & Bader, M. A unique central tryptophan hydroxylase isoform. *Biochem. Pharmacol.* **66**, 1673–1680 (2003).
- Thor, K. B. & Helke, C. J. Serotonin- and substance P-containing projections to the nucleus tractus solitarius of the rat. *J. Comp. Neurol.* **265**, 275–293 (1987).
- Abizaid, A. & Horvath, T. L. Brain circuits regulating energy homeostasis. *Regul. Pept.* **149**, 3–10 (2008).
- Grill, H. J. Distributed neural control of energy balance: contributions from hindbrain and hypothalamus. *Obesity (Silver Spring)* **14** (Suppl 5), 216S–221S (2006).
- Berthoud, H. R. & Morrison, C. The brain, appetite, and obesity. *Annu. Rev. Psychol.* **59**, 55–92 (2008).
- Heisler, L. K. *et al.* Serotonin reciprocally regulates melanocortin neurons to modulate food intake. *Neuron* **51**, 239–249 (2006).
- Xu, Y. *et al.* 5-HT_{2c}Rs expressed by pro-opiomelanocortin neurons regulate energy homeostasis. *Neuron* **60**, 582–589 (2008).
- Xu, Y. *et al.* A serotonin and melanocortin circuit mediates D-fenfluramine anorexia. *J. Neurosci.* **30**, 14630–14634 (2010).
- Takase, L. F. & Nogueira, M. I. Patterns of fos activation in rat raphe nuclei during feeding behavior. *Brain Res.* **1200**, 10–18 (2008).
- Fulwiler, C. E. & Saper, C. B. Subnuclear organization of the efferent connections of the parabrachial nucleus in the rat. *Brain Res.* **319**, 229–259 (1984).
- Rask-Andersen, M., Olszewski, P. K., Levine, A. S. & Schioth, H. B. Molecular mechanisms underlying anorexia nervosa: focus on human gene association studies and systems controlling food intake. *Brain Res. Brain Res. Rev.* **62**, 147–164 (2010).
- Kaye, W. Neurobiology of anorexia and bulimia nervosa. *Physiol. Behav.* **94**, 121–135 (2008).
- Tsien, J. Z., Huerta, P. T. & Tonegawa, S. The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell* **87**, 1327–1338 (1996).
- Hnasko, T. S. *et al.* Vesicular glutamate transport promotes dopamine storage and glutamate corelease *in vivo*. *Neuron* **65**, 643–656 (2010).
- Kremer, E. J., Boutin, S., Chillon, M. & Danos, O. Canine adenovirus vectors: an alternative for adenovirus-mediated gene transfer. *J. Virol.* **74**, 505–512 (2000).
- Kapli, M. G. *et al.* Long-term gene expression and phenotypic correction using adeno-associated virus vectors in the mammalian brain. *Nature Genet.* **8**, 148–154 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Froelick, J. Wang and K. Battani for help with histology; A. Rainwater for help with mouse breeding; A. Quintana for propagating CAV2-Cre virus and preparing AAV1-CreGFP virus; and A. Guler and M. Carter for helpful comments on the manuscript. This work was supported in part by National Institutes of Health grant DA024908 to R.D.P.

Author Contributions Q.W. and R.D.P. designed the research. Q.W. performed experiments and analysed the data. M.C. provided the conditional *Tph2* mouse line. R.D.P. and Q.W. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.D.P. (palmiter@uw.edu).

METHODS

Animal maintenance and neuron ablation. Mice were housed in a temperature- and humidity-controlled facility with a 12-h light cycle. All animal care and experimental procedures were approved by the Institutional Animal Care and Use Committee at the University of Washington. In compliance with our approved protocol, all experiments were terminated if the body weight of mice fell to 80% of their original body weight or they seemed moribund. *Agrp*^{DTR/+} mice, *Grin1*^{lox/lox} mice and *Slc17a6*^{lox/lox} mice were generated and genetically identified by PCR of tail DNA as described previously^{4,27,28}. A conditional *Tph2* targeting construct was prepared by flanking the first exon with *loxP* sites along with a *flr*-flanked *SV-Neo* gene in the first intron. The construct was electroporated in G4 ES cells, and correctly targeted clones were identified by Southern blotting. After removal of the *SV-Neo* gene by breeding with a mouse expressing FLP recombinase, heterozygotes were bred to generate *Tph2*^{lox/lox} mice that were used for viral injection. Details are available from the authors on request.

All except *Tph2*^{lox/lox} mice were on the C57Bl/6 background (more than nine generations backcrossed); *Tph2*^{lox/lox} mice were on a mixed 129/Sv × C57Bl/6 background. *Agrp*^{DTR/DTR} male mice were bred with *Slc17a6*^{lox/lox} female mice to generate *Agrp*^{DTR/+}; *Slc17a6*^{lox/+} mice, which were further bred to each other to create *Agrp*^{DTR/+}; *Slc17a6*^{lox/lox} mice and *Agrp*^{DTR/+}; *Slc17a6*^{lox/+} control mice. A similar breeding strategy was adopted when generating *Agrp*^{DTR/+}; *Tph2*^{lox/lox} mice and *Agrp*^{DTR/+}; *Grin1*^{lox/lox} mice. Mice were group housed with a standard chow diet (LabDiet 5053) and water provided *ad libitum* until the beginning of the experiments. All experiments were performed with male mice about 8 weeks old. To ablate AgRP neurons, mice carrying the *Agrp*^{DTR} allele were twice injected intramuscularly with DT (50 µg kg⁻¹; List Biological Laboratories, Campbell); the injections were 2 days apart⁴. The extent of ablation (more than 95%) was determined by immunohistochemistry⁴⁻⁷.

Viral injections. Production of adeno-associated virus 1 (AAV1-*CreGFP*) and canine adenovirus 2 (CAV2-*Cre*) followed the protocols described previously^{29,30}. For injection of CAV2-*Cre* or AAV1-*CreGFP* virus into the rostral NTS, mice were anaesthetized and virus (or PBS as the vehicle) was injected bilaterally or unilaterally (1 µl of roughly 10¹⁰ particles per µl each side) through a Hamilton syringe (size 5 µl; Hamilton, Reno), using stereotactic coordinates ±0.8 mm (x axis), -7.1 mm (y axis) and -4.3 mm (z axis). Similarly, AAV1-*CreGFP* virus (or PBS as the vehicle) was injected into the PBN bilaterally or unilaterally, using stereotactic coordinates ±1.0 mm (x axis), -5.3 mm (y axis) and -3.3 mm (z axis). Brain samples from all mice were collected at the end of the behavioural experiment and processed for immunohistological analysis. For all viral injection experiments, only a fraction (indicated in figure legends) of the DT-treated mice survived AgRP-neuron ablation; subsequent evaluation of *CreGFP* expression revealed that failure to rescue was associated with poor placement or inadequate viral transduction.

Drug treatments. Alzet 14-day minipumps (model 1002; Durect, Cupertino) loaded with 100 µl of ondansetron (6 mg ml⁻¹ in saline; Sigma-Aldrich, St Louis) were implanted subcutaneously on the back of anaesthetized mice 4 days before DT treatment. These minipumps dispense 0.25 µl h⁻¹. Alternatively, cannulas (28 gauge; Plastics One, Roanoke) were placed into fourth ventricles

under anaesthesia and the subcutaneous minipumps, which were loaded with ondansetron (0.6 mg ml⁻¹ in saline), were connected to the cannulas by tubing (PE60; Stoelting, Wood Dale) that was threaded under the skin to help prevent the mice from dislodging it. For some experiments, the minipumps (ondansetron, 0.6 mg ml⁻¹) were connected to bilateral cannulas (28 gauge; Plastics One) directed to specific brain regions by using the following stereotactic coordinates: PBN, ±1 mm (x axis), -5.3 mm (y axis) and -3.3 mm (z axis); NTS, ±0.8 mm (x axis), -7.1 mm (y axis) and -4.3 mm (z axis). The patency and placement of the bilateral minipump were verified at the end of each experiment, and brain samples were processed for immunohistological analysis.

Food intake and body weight measurements. For feeding assays, mice were transferred to BioDAQ Food and Water Intake Monitor (Research Diets, New Brunswick) supplied with water and low-fat (3.85 kcal ml⁻¹) chow diet (D12450B; Research Diets). The mice were allowed to acclimatize for 3 days before initiation of each experiment and data collection. Body weight and total food intake were recorded every 24 h. Feeding and drinking activity was recorded in accordance with the manufacturer's suggested protocol.

Immunohistochemistry. Mice were killed by CO₂ asphyxiation and perfused transcardially with ice-cold PBS buffer containing 4% paraformaldehyde. Brains were dissected and postfixed overnight at 4 °C in the fixation buffer. Free-floating brain sections (25 µm) were washed three times in PBS containing 0.1% Triton X-100 (PBST buffer) solution (15 min each wash) and then blocked with 3% normal donkey serum in PBST for 2–3 h at about 23 °C. Rabbit anti-AgRP (dilution 1:1,500; Phoenix Pharmaceuticals, Belmont), rabbit anti-Fos (dilution 1:1,000; Millipore, Temecula), monoclonal anti-tryptophan hydroxylase (dilution 1:1,500; Sigma-Aldrich) and rabbit anti-dopa decarboxylase (equivalent to AADC; dilution 1:500; Millipore) were applied to the sections for incubation overnight at 4 °C, followed by three 15-min rinses in PBST. Finally, sections were incubated in Cy2- or Cy3-labelled secondary antibody (dilution 1:300; Jackson Immunolaboratory, West Grove) before visualization. Images were captured with a digital camera mounted on a Leica TCS SP1 confocal microscope (Leica Microsystems); all paired photos were obtained through the same system settings. For each group of mice, at least eight sections from four different mice were analysed.

Data analyses. Quantification of *Tph2*-positive and AADC-positive cells was performed with the NIH Image software (National Institutes of Health). Anatomical correlations of brain sections and delineation of individual nuclei were determined by comparing landmarks of Nissl staining images with those given in the stereotactic atlas. From the anatomically matched sections, a region of interest of the same size was further defined. Meanwhile, an optimized threshold that can discern round nuclei from partly stained ones as well as background noise was preset for all measurements. For all experiments only those mice with correct placement of cannula or viral injections were compared with the control group. Unless otherwise stated, data sets collected from all experiments were analysed by one-way ANOVA followed by the Student–Newman–Keuls method for statistical significance; results were plotted as means ± s.e.m. Post-hoc analysis was performed when group differences were significant by ANOVA at *P* < 0.05.

Chromatin-modifying enzymes as modulators of reprogramming

Tamer T. Onder^{1,2,3,4}, Nergis Kara⁵, Anne Cherry^{1,2,3,4}, Amit U. Sinha^{6,7}, Nan Zhu^{3,6,7}, Kathrin M. Bernt^{3,6,7}, Patrick Cahan^{1,2,3,4}, B. Ogan Mancarci⁸, Juli Unternaehrer^{1,2,3,4}, Piyush B. Gupta^{9,10}, Eric S. Lander^{9,11,12}, Scott A. Armstrong^{3,6,7} & George Q. Daley^{1,2,3,4,6,13,14}

Generation of induced pluripotent stem cells (iPSCs) by somatic cell reprogramming involves global epigenetic remodelling¹. Whereas several proteins are known to regulate chromatin marks associated with the distinct epigenetic states of cells before and after reprogramming^{2,3}, the role of specific chromatin-modifying enzymes in reprogramming remains to be determined. To address how chromatin-modifying proteins influence reprogramming, we used short hairpin RNAs (shRNAs) to target genes in DNA and histone methylation pathways, and identified positive and negative modulators of iPSC generation. Whereas inhibition of the core components of the polycomb repressive complex 1 and 2, including the histone 3 lysine 27 methyltransferase EZH2, reduced reprogramming efficiency, suppression of SUV39H1, YY1 and DOT1L enhanced reprogramming. Specifically, inhibition of the H3K79 histone methyltransferase DOT1L by shRNA or a small molecule accelerated reprogramming, significantly increased the yield of iPSC colonies, and substituted for KLF4 and c-Myc (also known as MYC). Inhibition of DOT1L early in the reprogramming process is associated with a marked increase in two alternative factors, NANOG and LIN28, which play essential functional roles in the enhancement of reprogramming. Genome-wide analysis of H3K79me2 distribution revealed that fibroblast-specific genes associated with the epithelial to mesenchymal transition lose H3K79me2 in the initial phases of reprogramming. DOT1L inhibition facilitates the loss of this mark from genes that are fated to be repressed in the pluripotent state. These findings implicate specific chromatin-modifying enzymes as barriers to or facilitators of reprogramming, and demonstrate how modulation of chromatin-modifying enzymes can be exploited to more efficiently generate iPSCs with fewer exogenous transcription factors.

To examine the influence of chromatin modifiers on somatic cell reprogramming, we used a loss-of-function approach to interrogate the role of 22 select genes in DNA and histone methylation pathways. We tested a pool of three hairpins for each of 22 target genes and observed knockdown efficiencies of >60% for 21 out of 22 targets (Supplementary Fig. 1). We infected fibroblasts differentiated from the H1 human embryonic stem cell (ESC) line (dH1fs) with shRNA pools, transduced them with reprogramming vectors expressing OCT4 (also known as POU5F1), SOX2, KLF4 and c-Myc (OSKM), and identified the resulting iPSCs by Tra-1-60 staining (Fig. 1a)⁴. Eight shRNA pools reduced reprogramming efficiency (Fig. 1b). Among the target genes were *OCT4* (included as a control), and *EHMT1* and *SETDB1*, two H3K9 methyltransferases whose histone mark is associated with transcriptional repression. The remaining five shRNA

pools targeted components of polycomb repressive complexes (PRC), major mediators of gene silencing and heterochromatin formation⁵. Inhibition of PRC1 (BMI1, RING1) and PRC2 components (EZH2, EED, SUZ12) significantly decreased reprogramming efficiency while having negligible effects on cell proliferation (Fig. 1c and Supplementary Fig. 2). This finding is of particular significance given that EZH2 is necessary for fusion-based reprogramming⁶ and highlights the importance of transcriptional silencing of the somatic cell gene expression program during generation of iPSCs.

In contrast to genes whose functions seem to be required for reprogramming, inhibition of three genes enhanced reprogramming: *YY1*, *SUV39H1* and *DOT1L* (Fig. 1b, d). YY1 is a context-dependent transcriptional activator or repressor⁷, whereas SUV39H1 is a histone H3K9 methyltransferase implicated in heterochromatin formation⁸. Interestingly, enzymes that modify H3K9 were associated with both inhibition and enhancement of reprogramming, which suggested that unravelling the mechanisms for their effects might be challenging. Thus, we focused on DOT1L, a histone H3 lysine 79 methyltransferase that has not previously been studied in the context of reprogramming⁹. We used two hairpin vectors that resulted in the most significant downregulation of DOT1L and concomitant decrease in global H3K79 methylation levels (Supplementary Fig. 3a, b). Fibroblasts expressing *DOT1L* shRNAs formed significantly more iPSC colonies when tested separately or in a context where they were fluorescently labelled and co-mixed with control cells (Fig. 2a and Supplementary Fig. 4). This enhanced reprogramming phenotype could be reversed by overexpressing an shRNA-resistant wild-type DOT1L, but not a catalytically inactive DOT1L, indicating that inhibition of catalytic activity of DOT1L is key to enhance reprogramming¹⁰ (Fig. 2a). Our findings with dH1fs were applicable to other human fibroblasts, as IMR-90 and MRC-5 cells also showed threefold and sixfold increases in reprogramming efficiency, respectively, upon DOT1L suppression (Supplementary Fig. 5). To validate our findings independently of shRNA-mediated knockdown, we used a recently discovered small molecule inhibitor of DOT1L catalytic activity. EPZ004777 (ref. 11, referred to as iDot1L) abrogated H3K79 methylation at concentrations ranging from 1 μ M to 10 μ M and increased reprogramming efficiency three- to fourfold (Fig. 2b and Supplementary Fig. 6a, b). Combination of inhibitor treatment with DOT1L knockdown did not further increase reprogramming efficiency, reinforcing our previous observation that inhibition of the catalytic activity of DOT1L is key to reprogramming (Supplementary Fig. 6c). iPSCs generated through DOT1L inhibition showed characteristic ESC morphology, immunoreactivity for SSEA4, SSEA3, Tra-1-81, OCT4 and NANOG,

¹Stem Cell Transplantation Program, Division of Pediatric Hematology and Oncology, Manton Center for Orphan Disease Research, Children's Hospital Boston and Dana Farber Cancer Institute, Boston, Massachusetts, 02115, USA. ²Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, 02115, USA. ³Harvard Stem Cell Institute, Cambridge, Massachusetts, 02138, USA. ⁴Stem Cell Program, Children's Hospital Boston, Boston, Massachusetts, 02115, USA. ⁵German Cancer Research Center, Heidelberg, 69120, Germany. ⁶Division of Hematology/Oncology, Children's Hospital, Harvard Medical School, Boston, Massachusetts, 02115, USA. ⁷Department of Pediatric Oncology, Harvard Medical School, Boston, Massachusetts, 02115, USA. ⁸Department of Molecular Biology and Genetics, Bilkent University, Ankara, 06800, Turkey. ⁹Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02142, USA.

¹⁰Whitehead Institute for Biomedical Research, Cambridge, Massachusetts, 02142, USA. ¹¹The Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, 02142, USA. ¹²Department of Systems Biology, Harvard Medical School, Boston, Massachusetts, 02115, USA. ¹³Division of Hematology, Brigham and Women's Hospital, Boston, Massachusetts, 02115, USA.

¹⁴Howard Hughes Medical Institute, Chevy Chase, Maryland, 20815, USA.

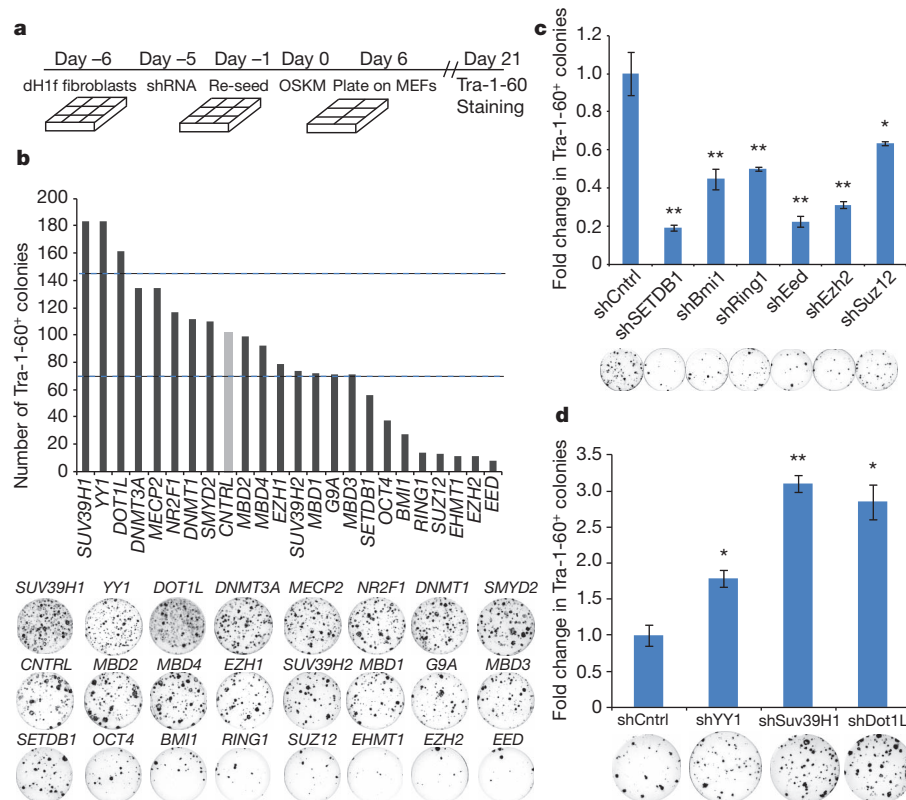


Figure 1 | Screening for inhibitors and enhancers of reprogramming.

a, Timeline of shRNA infection and iPSC generation. **b**, Number of Tra-1-60⁺ colonies 21 days after OSKM transduction of 25,000 dH1f cells previously infected with pools of shRNAs against the indicated genes. Representative Tra-1-60-stained reprogramming wells are shown. The dotted lines indicates 3 standard deviations from the mean number of colonies in control wells. **c**, Validation of primary screen hits that decrease reprogramming efficiency.

Fold change in Tra-1-60⁺ iPSC colonies relative to control cells. * $P < 0.05$, ** $P < 0.01$ compared to control shRNA-expressing fibroblasts ($n = 4$; error bars, \pm s.e.m.). Representative Tra-1-60-stained wells are shown. **d**, Validation of primary screen hits that increase reprogramming efficiency. Fold change in Tra-1-60⁺ iPSC colonies relative to control cells. * $P < 0.05$, ** $P < 0.01$ compared to control shRNA-expressing fibroblasts ($n = 4$; error bars, \pm s.e.m.). Representative Tra-1-60-stained wells are shown.

and differentiated into all three embryonic germ layers *in vitro* and in teratomas (Supplementary Fig. 7a–c). Therefore, iPSCs generated following DOT1L inhibition display all of the hallmarks of pluripotency.

We next assessed DOT1L inhibition in murine reprogramming. iDot1L treatment led to threefold enhancement of reprogramming of mouse embryonic fibroblasts carrying an OCT4-GFP (green fluorescent protein) reporter gene (OCT4-GFP MEFs; Fig. 2c). Reprogramming of tail-tip fibroblasts (TTFs) derived from a conditional knockout *DOT1L* mouse strain yielded significantly more iPSC colonies upon deletion of *DOT1L*¹² (Supplementary Fig. 8a). Cre-mediated excision of both floxed *DOT1L* alleles in iPSC clones derived from homozygous TTFs was confirmed by genomic PCR (Supplementary Fig. 8b). DOT1L inhibition also increased reprogramming efficiency of MEFs and peripheral blood cells derived from an inducible secondary iPSC mouse strain¹³ (Supplementary Fig. 8c, d). Taken together, these results demonstrate that DOT1L inhibition enhances reprogramming of both mouse and human cells.

We next examined the cellular mechanisms by which DOT1L inhibition promotes reprogramming. DOT1L inhibition affected neither retroviral transgene expression nor cellular proliferation (Supplementary Fig. 9a–c). Although previous studies indicated that *DOT1L*-null cells have increased apoptosis and accumulation of cells in G2 phase⁹, we failed to observe a significant increase in apoptosis or change in the cell cycle profile of DOT1L-inhibited fibroblasts (Supplementary Fig. 9d, e). In human iPSC clones derived from shDot1L fibroblasts, DOT1L inhibition was no longer evident, reflecting the known silencing of retroviruses that occurs during reprogramming (Supplementary Fig. 10a). Quantitative PCR (qPCR) analysis

revealed that the silencing occurred by day 15 after OSKM transduction (Supplementary Fig. 10b, c). To define the crucial time window for DOT1L inhibition, we treated fibroblasts with iDot1L at 1-week intervals during reprogramming. iDot1L treatment in either the first or second week was sufficient to enhance reprogramming, whereas treatment in the third week or a 5-day pretreatment had no effect (Supplementary Fig. 10d, e). Immunofluorescence analysis revealed significantly greater numbers of Tra-1-60-positive cell clusters on day 10 and day 14 in shDot1L cultures (Supplementary Fig. 11a, b), indicating that the emergence of iPSCs is accelerated upon DOT1L inhibition. When we extended the reprogramming experiments by 10 more days, shDot1L cells still yielded more iPSC colonies than controls (Supplementary Fig. 11c). Taken together, these findings indicate that DOT1L inhibition acts in early to middle stages to accelerate and increase the efficiency of the reprogramming process.

To assess whether DOT1L inhibition could replace any of the reprogramming factors, we infected control and DOT1L-inhibited fibroblasts with three factors, omitting one factor at a time. In the absence of OCT4 or SOX2 no iPSC colonies emerged (Fig. 2d). When we omitted either KLF4 or c-Myc, DOT1L-inhibited fibroblasts gave rise to robust numbers of Tra-1-60-positive colonies, whereas control cells generated very few colonies, as reported previously⁴ (Fig. 2d–f and Supplementary Fig. 12a). Importantly, DOT1L-inhibited fibroblasts transduced with only OCT4 and SOX2 gave rise to Tra-1-60-positive colonies, whereas control fibroblasts did not (Fig. 2d–f). These two-factor iPSCs showed typical ESC morphology, silenced the reprogramming vectors and had all of the hallmarks of pluripotency as gauged by endogenous pluripotency factor expression and the ability to form all three embryonic germ layers *in vitro* and in teratomas (Supplementary

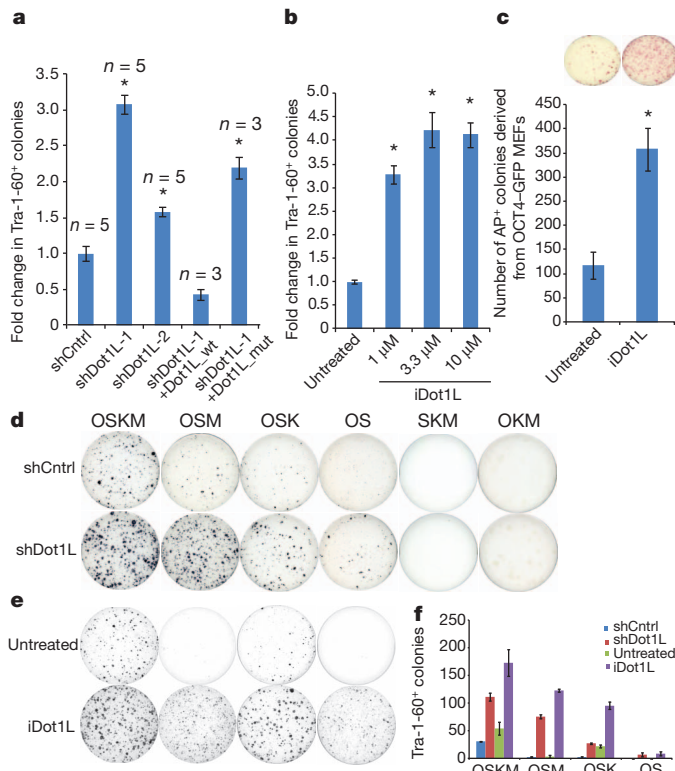


Figure 2 | DOT1L inhibition enhances reprogramming efficiency and substitutes for KLF4 and Myc. **a**, Fold change in the reprogramming efficiency of dH1f cells infected with two independent *DOT1L* shRNAs or co-infected with shRNA-1 and a vector expressing an shRNA-resistant wild-type or catalytically dead mutant *DOT1L*. Data correspond to the average and s.e.m.; $n =$ independent experiments. * $P < 0.01$ compared to control shRNA-expressing fibroblasts. **b**, Fold change in the reprogramming efficiency of dH1f cells treated with iDot1L at the indicated concentrations for 21 days. Data correspond to the mean \pm s.d.; $n = 3$. * $P < 0.001$ compared to untreated fibroblasts. **c**, Number of alkaline-phosphatase-positive (AP⁺) colonies derived from OSKM-transduced untreated or iDot1L-treated (10 μ M) OCT4-GFP MEFs. * $P < 0.001$ compared untreated MEFs ($n = 4$; error bars, \pm s.d.). Representative AP-stained wells are shown. **d**, Tra-1-60 stained of plates of shCntrl and shDot1L fibroblasts in the absence of each factor or both KLF4 and c-Myc. **e**, Tra-1-60-stained plates of untreated and iDot1L treated (3.3 μ M) fibroblasts in the absence of each factor or both KLF4 and c-Myc. **f**, Quantification of the Tra-1-60⁺ colonies in Fig. 2d, e representing mean and s.d. of two independent experiments done in triplicate.

Figs 7a–c and 12b). PCR on genomic DNA isolated from expanded colonies confirmed the absence of integrated KLF4 and c-Myc transgenes (Supplementary Fig. 12c). Thus, we were able to generate two-factor iPSCs either by suppression of *DOT1L* expression or chemical inhibition of its methyltransferase activity.

To gain insights into the molecular mechanisms of how *DOT1L* inhibition promotes reprogramming and replaces KLF4 we performed global gene-expression analyses on control and shDot1L fibroblasts before and 6 days after OSKM and OSM transduction, along with cells that were treated with iDot1L. Relatively few genes were differentially expressed in shDot1L cells on day 6 of reprogramming (22 up, 23 down; Supplementary Table 3). Inhibitor-treated cells showed broader gene expression changes (405 up and 175 down; Supplementary Table 3), presumably due to more complete inhibition of K79me2 levels (Fig. 3a). In the absence of KLF4, 94 genes were differentially upregulated in shDot1L cells; intersection of this set of genes with the set differentially upregulated in four-factor reprogramming of *DOT1L*-inhibited cells yielded only five common genes (Fig. 3a, b). We were particularly intrigued to find *NANOG* and *LIN28* upregulated in all three instances of *DOT1L* inhibition, because these two genes are part of the core pluripotency network of human ESCs^{14,15} and can reprogram human

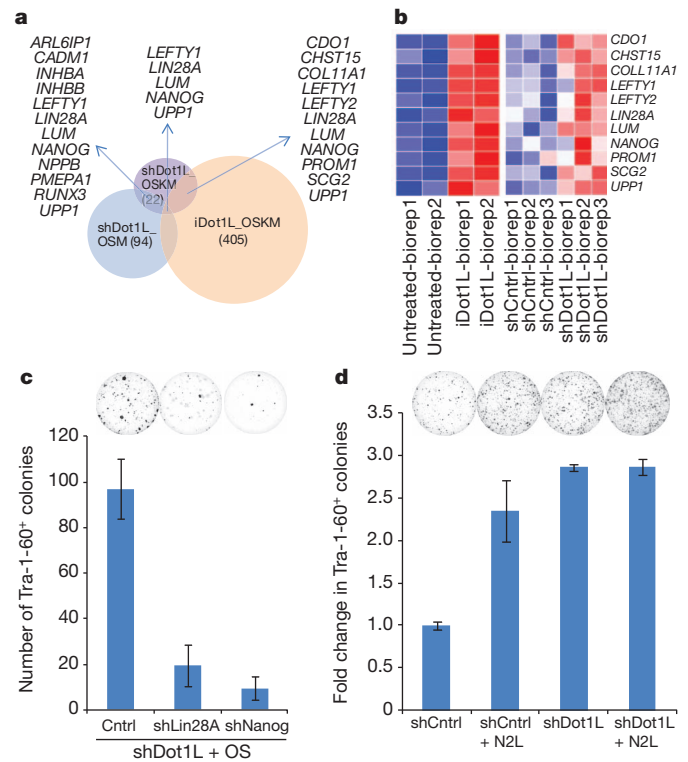


Figure 3 | NANOG and LIN28 are required for enhancement of reprogramming by *DOT1L* inhibition. **a**, Overlap of differentially upregulated genes in shDot1L cells 6 days post-OSKM and OSM transduction with the genes upregulated in OSKM-transduced iDot1L-treated cells. **b**, Heat maps showing differential expression levels of commonly upregulated genes in OSKM-transduced *DOT1L*-inhibited cells. **c**, Number of Tra-1-60⁺ iPSC colonies upon knockdown of Nanog or Lin28 in 2-factor reprogramming of shDot1L cells. Data represent mean and s.e.m. of 2 independent experiments done in triplicate. **d**, Fold-change in Tra-1-60⁺ iPSC colonies in 4-factor (OSKM) and 6-factor (OSKMNL) reprogramming of shCntrl and shDot1L fibroblasts. Data represent mean and s.e.m. of two independent experiments done in duplicate. Representative Tra-1-60-stained wells are shown above.

fibroblasts into iPSCs when used in combination with *OCT4* and *SOX2* (ref. 16).

We explored the possibility that *NANOG* and *LIN28* upregulation might account for the enhanced reprogramming observed following *DOT1L* inhibition, and validated their upregulation in shDot1L fibroblasts upon OSM or OS transduction (Supplementary Fig. 13a, b). Interestingly, at this early time point *REX1* (also known as *ZFP42*) and *DNMT3B*, two other well-characterized pluripotency genes, were not upregulated, indicating that *DOT1L* inhibition does not broadly upregulate the pluripotency network. Suppression of either Nanog or Lin28 abrogated the two-factor (OS) reprogramming of shDot1L fibroblasts, indicating the essential roles of *NANOG* and *LIN28* in this process (Fig. 3c and Supplementary Fig. 13c). *DOT1L* inhibition also led to increased *NANOG* expression in the context of *OCT4*, *SOX2* and *LIN28* (OSL) and *LIN28* expression in the context of *OCT4*, *SOX2* and *NANOG* (OSN) (Supplementary Fig. 14a). Furthermore, *DOT1L* inhibition significantly increased the efficiency of three-factor reprogramming in the context of OSN and OSL (Supplementary Fig. 14b). Finally, inclusion of *NANOG* and *LIN28* in the OSKM reprogramming cocktail did not confer any additional enhancement to shDot1L cells (Fig. 4d and Supplementary Fig. 14c). Taken together, these data implicate *NANOG* and *LIN28* in the enhancement of reprogramming and replacement of KLF4 and c-Myc with *DOT1L* inhibition.

To gain insight into the genome-wide chromatin changes that are facilitated by *DOT1L* inhibition, we performed chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) for H3K79me2 and H3K27me3 in human ESCs as well as fibroblasts undergoing

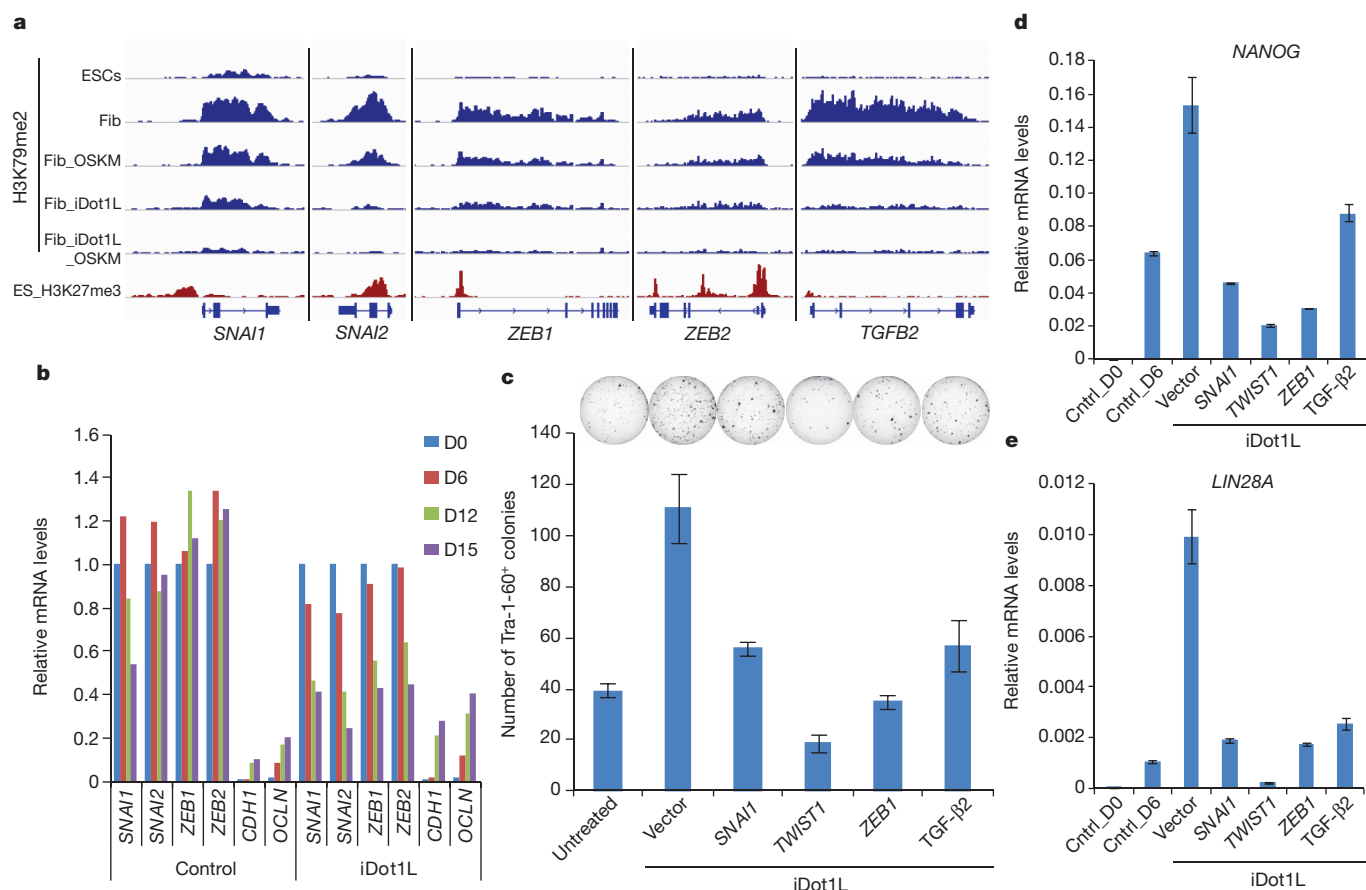


Figure 4 | Genome-wide analysis of H3K79me2 marks during reprogramming. **a**, H3K79me2 ChIP-seq tracks (blue) for select EMT-associated genes in fibroblasts (Fib) and ESCs along with the corresponding H3K27me3 tracks in ESCs (red). **b**, Expression of EMT-associated transcription factors (EMT-TF) and epithelial genes in control and iDot1L-treated fibroblasts at the indicated time points during reprogramming. qPCR was normalized to uninfected fibroblasts for EMT-TFs and H1 ESCs for *CDH1* and *OCLN*. **c**, Number of Tra-1-60⁺ colonies derived from untreated and iDot1L-treated (3.3 μ M) dH1f cells that are either infected with *SNAI1*,

TWIST1 or *ZEB1* expression vectors or treated with soluble TGF- β 2 (2 ng ml⁻¹) ($n = 3$; error bars, \pm s.d.). Representative Tra-1-60-stained wells are shown. **d**, qRT-PCR quantification of *NANOG* mRNA level on day 6 of OSKM-expressing untreated or iDot1L-treated (3.3 μ M) fibroblasts expressing the indicated EMT-factors. Expression levels were normalized to those observed in H1 ESCs. **e**, qRT-PCR quantification of *LIN28A* mRNA level on day 6 of OSKM-expressing untreated or iDot1L-treated (3.3 μ M) fibroblasts expressing the indicated EMT-factors. Expression levels were normalized to those observed in H1 ESCs.

reprogramming, with or without iDot1L treatment (Supplementary Fig. 15). In both ESCs and fibroblasts, H3K79me2 is positively associated with transcriptionally active genes and negatively associated with genes marked by H3K27me3 (Supplementary Fig. 16a–c). ESC-specific genes marked by H3K79me2 included pluripotency factors, a subset of their downstream targets, and genes involved in epithelial cell adhesion such as *CDH1* (E-cadherin) (280 genes; Supplementary Fig. 17a, b and Supplementary Tables 4, 5). In contrast, in fibroblasts, genes marked by H3K79me2 were significantly enriched in genes induced during the epithelial to mesenchymal transition (EMT) (377 genes; Supplementary Fig. 17a).

Among the 348 genes that showed reduced H3K79me2 6 days after OSKM expression, we likewise found a significant enrichment of gene sets associated with the induction of a mesenchymal state, including *SNAI2*, *TGFB2* and *TGFBRI* (Supplementary Fig. 18a)^{17,18}. Only a few of these genes showed decreased expression at day 6 (12 out of 348), but the vast majority of them lacked this mark in the pluripotent state (272 out of the 348 devoid of H3K79me2 in ESCs), suggesting they were destined for transcriptional silencing during reprogramming. This finding prompted us to ask whether DOT1L inhibition results in the removal of H3K79me2 from such fibroblast-specific, EMT-associated genes. Upon DOT1L inhibitor treatment, H3K79me2 levels were reduced on almost all loci, with the exception of a subset comprised mostly of housekeeping genes that also had high levels of H3K79me2 in ESCs (Supplementary Fig. 19a). Strikingly, the genes

that lost proportionally the most H3K79me2 in inhibitor-treated fibroblasts during reprogramming (eightfold or more) were again highly enriched in genes induced in EMT (Supplementary Fig. 19b). Mesenchymal master regulators such as *SNAI1*, *SNAI2*, *ZEB1*, *ZEB2* and *TGFB2* were among these genes (Fig. 4a)¹⁹. In the presence of the DOT1L inhibitor, these regulators were more strongly repressed during reprogramming, whereas epithelial genes such as *CDH1* and *OCLN* were more robustly upregulated (Fig. 4b). The extinction of fibroblast gene expression was accompanied by increased deposition of the repressive H3K27me3 mark on the majority of fibroblast-specific regulators examined (Supplementary Fig. 20). In contrast, H3K27me3 was depleted to a greater extent on *SOX2* and E-cadherin promoters, reflecting their activation during reprogramming. Finally, the H3K27me3 status of master regulators of other lineages, such as *OLIG2*, *MYOD1*, *NKX2-1* and *GATA4*, remained unchanged upon DOT1L inhibitor treatment, indicating that the deposition of H3K27me3 was specific to fibroblast-specific regulators.

To test the functional importance of downregulation of mesenchymal regulators in the iDot1L-mediated enhancement of reprogramming, we overexpressed *TWIST1*, *SNAI1* and *ZEB1* or added soluble TGF- β 2 to cells undergoing reprogramming in the presence of the DOT1L inhibitor. All of these perturbations significantly counteracted the enhancement observed with DOT1L inhibition (Fig. 4c). Interestingly, expression of these factors also abrogated the iDot1L-mediated upregulation of *NANOG* and *LIN28*, suggesting that the effect of DOT1L

inhibition on these two pluripotency genes is likely to be indirect (Fig. 4d, e). Conversely, we tested whether destabilization of the mesenchymal state by inhibition of TGF- β signalling would be redundant with DOT1L inhibition. A small molecule inhibitor of TGF- β signalling (SB431542) increased reprogramming efficiency, but in combination with the DOT1L inhibitor, showed no significant further increase in iPSC colonies (Supplementary Fig. 21). Taken together these data indicate that in fibroblasts, downregulation of the mesenchymal gene expression program is critical to enhancement of reprogramming by DOT1L inhibition.

Our loss-of-function survey indicates that chromatin-modifying enzymes play critical roles for both reactivating silenced loci as well as reinstating closed domains of heterochromatin during the global epigenetic remodelling of differentiated cells to pluripotency, thus implicating specific enzymes as facilitators or barriers to cell fate transitions. DOT1L inhibition seems to enhance reprogramming at least in part by facilitating loss of H3K79me2 from fibroblast genes whose silencing is required for reprogramming (Supplementary Fig. 22). Interestingly, KLF4, which can be replaced by DOT1L inhibition, has been shown to facilitate a mesenchymal to epithelial transition (MET) by inducing E-cadherin expression²⁰. Persistent H3K79me2 at the fibroblast master regulators during the initial phases of reprogramming seems to prevent shutdown of these genes, thus hindering the acquisition of an epithelial phenotype concomitant with delayed activation of NANOG and LIN28. In this regard H3K79me2 acts as a barrier to efficient repression of the somatic program by the reprogramming factors. This notion is consistent with the role of Dot1 in yeast, where it antagonizes gene repression²¹. As reprogramming of blood cells is also enhanced by DOT1L inhibition, we speculate that DOT1L inhibition may enhance reprogramming in a broad range of cell types by facilitating the silencing of lineage-specific programs of gene expression. Finally, our results also demonstrate that specific chromatin modifiers can be modulated to generate iPSCs more efficiently and with fewer exogenously introduced transcription factors.

METHODS SUMMARY

shRNAs were designed using the RNAi Codex²². 97-mer oligonucleotides (Supplementary Table 1) were PCR-amplified and cloned into the MSCV-PM²³ vector. Reprogramming assays were carried out with either retroviral⁴ or lentiviral¹⁶ reprogramming vectors. dH1f cells were previously described⁴. For gene expression analyses, total RNA was extracted from two or three independent culture plates for each condition and transcriptional profiling was performed using Affymetrix U133A microarrays. ChIP-seq was performed as described with slight modifications¹².

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 May 2011; accepted 16 February 2012.

Published online 4 March; corrected 28 March 2012 (see full-text HTML version for details).

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
2. Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479–491 (2010).
3. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).

4. Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
5. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
6. Pereira, C. F. *et al.* ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* **6**, 547–556 (2010).
7. Shi, Y. *et al.* Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* **67**, 377–388 (1991).
8. Schotta, G., Ebert, A. & Reuter, G. S. U. (VAR)3–9 is a conserved key function in heterochromatic gene silencing. *Genetica* **117**, 149–158 (2003).
9. Jones, B. *et al.* The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* **4**, e1000190 (2008).
10. Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
11. Daigle, S. R. *et al.* Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53–65 (2011).
12. Bernt, K. M. *et al.* MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).
13. Carey, B. W. *et al.* Single-gene transgenic mouse strains for reprogramming adult somatic cells. *Nature Methods* **7**, 56–59 (2010).
14. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
15. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
16. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
17. Charafe-Jauffret, E. *et al.* Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene* **25**, 2273–2284 (2006).
18. Onder, T. T. *et al.* Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res.* **68**, 3645–3654 (2008).
19. Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl Acad. Sci. USA* **107**, 15449–15454 (2010).
20. Samavarchi-Tehrani, P. *et al.* Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
21. Stulemeijer, I. J. *et al.* Dot1 binding induces chromatin rearrangements by histone methylation-dependent and -independent mechanisms. *Epigenetics chromatin* **4**, 2 (2011).
22. Olson, A. *et al.* RNAi Codex: a portal/database for short-hairpin RNA (shRNA) gene-silencing constructs. *Nucleic Acids Res.* **34**, D153–D157 (2006).
23. Schlachbach, M. R. *et al.* Cancer proliferation gene discovery through functional genomics. *Science* **319**, 620–624 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Hu and S. J. Elledge for providing the MSCV-PM vector, K. Ng and M. W. Lensch for teratoma injections and assessment and S. Loewer for discussions. We also thank E. Olhava and Epizyme Inc. for synthesizing and providing the DOT1L inhibitor, EPZ004777. G.Q.D. is an investigator of the Howard Hughes Medical Institute. Research was funded by grants from the US National Institutes of Health (NIH) to S.A.A. (CA140575) and G.Q.D., and the CHB Stem Cell Program.

Author Contributions T.T.O. performed project planning, experimental work, data interpretation and preparation of the manuscript. N.K., A.C., N.Z., J.U. and B.O.M. performed experimental work. P.C. and A.U.S. participated in data analysis. K.M.B. and S.A.A. provided critical materials and participated in the preparation of the manuscript. P.B.G. and E.S.L. participated in data acquisition, data interpretation and preparation of the manuscript. G.Q.D. supervised research and participated in project planning, data interpretation and preparation of the manuscript.

Author Information The microarray and ChIP-seq data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) and are accessible through GEO Series accession numbers GSE29253 and GSE35791. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.Q.D. (George.Daley@childrens.harvard.edu).

METHODS

shRNA cloning. shRNAs were designed using the RNAi Codex²². 97-mer oligos (Supplementary Table 1) were amplified with the following primer pair: Forward: GATGGCTGCTCGAGAAGGTATATTGCTGTTGACAGTGAAGC, reverse: GTCTAGAGGAATTCGAGGAGCAGTAGGC. PCR products were gel-purified, digested with EcoRI and XhoI and ligated into the MSCV-PM vector. Clones were verified by sequencing. shRNA targeting the firefly luciferase was used as a control²³. NANOG shRNA was previously described²⁴.

Production of viral supernatants. 293T cells were plated at a density of 2.5×10^6 cells per 10-cm dish. The next day, cells were transfected with 2.5 µg viral vector, 2.25 µg Gag-Pol vector and 0.25 µg VSV-G plasmid using 20 µl Fugene 6 (Roche Applied Science) in 400 µl DMEM per plate. Supernatant was collected 48 h and 72 h post-transfection and filtered through 45-µm pore size filters. For concentration, viral supernatants were mixed with PEG3350 solution (Sigma P3640, dissolved in PBS, 10% final concentration) and left overnight at 4 °C. The next day, supernatants were centrifuged at 2,500 r.p.m. for 20 min, and the pellets were re-suspended in PBS. Titering was performed on 293Ts. For shRNA infections, 500 µl of unconcentrated viral supernatant was used to infect 25,000 cells in the presence of $10 \mu\text{g ml}^{-1}$ protamine sulphate. For fluorescent labelling of dH1fs, we used lentiviruses PRRL-GFP (Addgene catalogue no. 12252) and FUDGW-Tomato (Addgene catalogue no. 22771).

Reprogramming assays. dH1f cells were first infected with shRNA viruses at high multiplicity of infection (m.o.i.) to ensure all cells received at least one vector (gauged by puromycin resistance of parallel infected wells). 25,000 shRNA-infected dH1f cells were then plated per well in 12-well plates and infected overnight with either retroviral (m.o.i. 2.5×10^5) or lentiviral (Addgene catalogue no. 21162, 21164; 100–200 µl supernatant)²⁶ reprogramming factors. For two-factor reprogramming, OCT4 and SOX2 viruses were used at an m.o.i. of 5. Six days later, cells were trypsinized and re-plated 1:4 or 1:6 onto six-well plates. Medium was changed to hES medium daily until day 21 when plates were fixed. Small molecule inhibitor of DOT1L, EPZ004777 (a gift from Epizyme, Inc.) was dissolved in DMSO as a 10 mM stock and was added at the indicated concentrations. For DOT1L rescue experiments, an MSCV-based retroviral vector encoding human DOT1L with or without mutations in the SAM-binding site (gifts of Y. Zhang) were mutagenized at the shRNA target site using a QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). In certain experiments, NANOG and LIN28 expression was achieved using lentiviruses (Addgene catalogue no. 21163). IMR-90 and MRC5 human diploid fibroblasts were purchased from ATCC and 50,000 cells were used in reprogramming experiments.

SB431542 (Stemgent) was used at a final concentration of 2 µM. TGF-β2 (R&D Systems) was added daily at a concentration of 2 ng ml^{-1} . Twist1 (Addgene catalogue no. 1783), Snai1 (Addgene catalogue no. 23347) and Zeb1 (a gift of R. A. Weinberg) were overexpressed using retro- or lentiviruses. Statistical analysis was performed using a Student's *t*-test.

Microarray analysis. Total RNA was extracted from two or three independent culture plates for each condition with an RNeasy Mini kit (Qiagen). Synthesis of complementary RNA from total RNA and hybridization/scanning of microarrays were performed with Affymetrix GeneChip products (HGU133A) as described in the GeneChip manual. Normalization of the raw gene expression data, quality control checks and subsequent analyses were done with the open-source R-project statistical software (<http://www.r-project.org/>) together with Bioconductor packages. Raw data files (.CEL) were converted into probe set values by RMA (robust multi-array averaging) normalization. Genes were selected at a threshold of log ratio 0.4. The microarray data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE29253.

SYBR-Green real-time RT-PCR. Total RNA was extracted using an RNeasy Mini kit coupled with an RNase-free DNase set (Qiagen) and reverse transcribed with Hexanucleotide Mix (Roche). The resulting complementary DNAs were used for PCR using SYBR-Green Master PCR mix (Applied Biosystems) in triplicates. All quantifications were normalized to an endogenous β-actin control. The relative quantification value for each target gene compared to the calibrator for that target is expressed as $2^{-(C_t - C_c)}$ (C_t and C_c are the mean threshold cycle differences after normalizing to β-actin). List of primers can be found in Supplementary Table 2.

Immunostaining. Immunostaining of reprogramming plates was performed as described²⁷. Briefly, cells were fixed with 4% paraformaldehyde and stained with biotin-anti-Tra-1-60 (eBioscience, catalogue no. 13-8863-82, 1:250) and streptavidin horseradish peroxidase (Biolegend, catalogue no. 405210, 1:500) diluted in PBS (3%), FCS (0.3%) Triton X-100. Staining was developed with the Vector labs DAB kit (catalogue no. SK-4100), and iPSC colonies quantified with ImageJ software.

For the characterization of shDOT1L-iPSC cells, we picked single colonies onto MEF-coated 96-well plates. The plates were fixed for 20 min with 4% paraformaldehyde/phosphate-buffered saline with calcium and magnesium (PBS (+/+)), washed

several times with PBS (+/+) and incubated overnight at 4 °C with primary antibody and Hoechst stain diluted in 3% donkey serum/3% BSA Fraction VII/0.01% Triton X-100/PBS (+/+); Hoechst, Invitrogen catalogue no. H3570 (1:20,000), Tra-1-81/A488 (BD catalogue no. 560174), SSEA-4/A647 (BD catalogue no. 560219), Tra-1-60/A647 (BD catalogue no. 560122), Nanog, rabbit polyclonal (Abcam catalogue no. ab21624), OCT4, rabbit polyclonal (Abcam catalogue no. ab19857). For Nanog and OCT4, donkey anti-rabbit IgG/A555 (Molecular Probes catalogue no. A31572) secondary antibody was used. After several washes with PBS (+/+), images were acquired using a BD Pathway 435 imager equipped with a $\times 10$ objective.

Teratoma formation assay. iPSCs grown on MEFs were harvested with Collagenase IV (1 mg ml^{-1} in DMEM/F12). Cell clumps from one six-well plate were resuspended in 50 µl DMEM/F12, 100 µl collagen I (Invitrogen catalogue no. A1064401) and 150 µl hESC-qualified Matrigel (BD Biosciences#354277). Cell clumps were then injected into the hind limb femoral muscles (100 µl suspension per leg) of Rag2^{γ/c} mice. After 6–8 weeks, teratomas were harvested and fixed in Bouin's solution overnight. Samples were then embedded in paraffin, and sections were stained with haematoxylin/eosin (Rodent Histopathology Core, Harvard Medical School).

Characterization of iPSC cells. Embryoid body differentiation was performed as described²⁸. To check for the presence of the reprogramming transgenes, genomic DNA was isolated using DNeasy Blood & Tissue Kit (Qiagen) and PCR was performed with specific primers to the endogenous or the viral transgenes⁴.

ChIP-sequencing. ChIP-seq was performed as described with slight modifications¹². 300,000 cells were fixed at room temperature in PBS 1% formalin (v/v) for 10 min with gentle agitation. Fixation was stopped by the addition of glycine (125 mM final concentration) and agitation for 5 min at room temperature. Fixed cells were washed twice in ice-cold PBS, resuspended in 100 µl of SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1). Chromatin was sheared by sonication to about 100–500 base pair fragments using a Bioruptor (diagenode) and diluted tenfold with dilution buffer (0.01% SDS, 1.1% Triton-X100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl). Antibodies against specific histone modifications were added to sonicated chromatin solution and incubated at 4 °C overnight with gentle agitation. The antibodies used were anti-H3K27me3 (Millipore 07-449) and anti-H3K79me2 (abcam 3594). Immune complexes were collected by incubation with 20 µl of protein A/G agarose beads (Millipore) for an hour at 4 °C with gentle agitation. Precipitates were washed sequentially with ice-cold low-salt wash (0.1% SDS, 1% Triton-X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 150 mM NaCl), high-salt wash (0.1% SDS, 1% Triton-X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 500 mM NaCl), LiCl wash (0.25 M LiCl, 1% IGEPAL CA-630, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris-HCl, pH 8.1) and TE wash (1 mM EDTA, 10 mM Tris-HCl, pH 8.1) for 5 min each at 4 °C with gentle agitation. Samples were centrifuged briefly in between washes to collect the beads. Immunoprecipitated DNA was eluted by incubating beads with 150 µl elution buffer (1%SDS, 0.1 M NaHCO₃) with gentle agitation for 15 min at room temperature. Elution was repeated once and eluates were combined, sodium chloride (final concentration of 0.2 M) was added to the eluate and eluates were incubated at 65 °C overnight to reverse crosslinking. DNA was purified using PCR purification spin columns (Qiagen). For ChIP sequencing, ChIP DNA libraries were made following Illumina ChIP-seq library preparation kit and subjected to Solexa sequencing (Illumina) at the Center for Cancer Computational Biology, Dana Faber Cancer Institute. Sequencing was performed on Illumina HiSeq2000. The reads were aligned to the human genome hg18 using Bowtie²⁹ and the reads that mapped to multiple locations in the genome were discarded. We quantified the histone modification level as the number of reads per million per kilobase in a window of interest. The window was 1 kb upstream to 1 kb downstream from the transcription start site (TSS) for H3K27me3 and 1 kb upstream to 2 kb downstream of the TSS for H3K79me2. To determine the significance of signal at a gene, an empirical background model was estimated. Genes that showed interesting pattern of histone methylation change were identified using iCanPlot (<http://www.icanplot.org>). Geneset Overlap Analysis was performed by finding the overlap of a set of genes of interest with the gene sets in the collections c2.all, c3.all and c5.all in MSigDB (total number of genesets in these collections is 5,562)³⁰. Hypergeometric test, with Bonferroni correction for multiple hypothesis testing, was performed to generate the *P* values associated with gene set overlap analysis. ChIP-Seq data have been deposited at the NCBI Gene Expression Omnibus with accession number GSE35791.

24. Zaehres, H. *et al.* High-efficiency RNA interference in human embryonic stem cells. *Stem Cells* **23**, 299–305 (2005).

25. Park, I.-H. *et al.* Generation of human-induced pluripotent stem cells. *Nature Protocols* **3**, 1180–1186 (2008).

26. Yu, J. *et al.* Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**, 797–801 (2009).

27. Chan, E. M. *et al.* Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nature Biotechnol.* **27**, 1033–1037 (2009).
28. Loewer, S. *et al.* Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature Genet.* **42**, 1113–1117 (2010).
29. Langmead, B. *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
30. Subramanian, A. *et al.* GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics* **23**, 3251–3253 (2007).

The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity

Jordi Barretina^{1,2,3,†*}, Giordano Caponigro^{4*}, Nicolas Stransky^{1*}, Kavitha Venkatesan^{4*}, Adam A. Margolin^{1†*}, Sungjoon Kim⁵, Christopher J. Wilson⁴, Joseph Lehar⁴, Gregory V. Kryukov¹, Dmitriy Sonkin⁴, Anupama Reddy⁴, Manway Liu⁴, Lauren Murray¹, Michael F. Berger^{1†}, John E. Monahan⁴, Paula Morais¹, Jodi Meltzer⁴, Adam Korejwa¹, Judit Jané-Valbuena^{1,2}, Felipa A. Mapa⁴, Joseph Thibault⁵, Eva Bric-Furlong⁴, Pichai Raman⁴, Aaron Shipway⁵, Ingo H. Engels⁵, Jill Cheng⁶, Guoying K. Yu⁶, Jianjun Yu⁶, Peter Aspesi Jr⁴, Melanie de Silva⁴, Kalpana Jagtap⁴, Michael D. Jones⁴, Li Wang⁴, Charles Hatton³, Emanuele Palescandolo³, Supriya Gupta¹, Scott Mahan¹, Carrie Sougnez¹, Robert C. Onofrio¹, Ted Liefeld¹, Laura MacConaill³, Wendy Winckler¹, Michael Reich¹, Nanxin Li⁵, Jill P. Mesirov¹, Stacey B. Gabriel¹, Gad Getz¹, Kristin Ardlie¹, Vivien Chan⁶, Vic E. Myer⁴, Barbara L. Weber⁴, Jeff Porter⁴, Markus Warmuth⁴, Peter Finan⁴, Jennifer L. Harris⁵, Matthew Meyerson^{1,2,3}, Todd R. Golub^{1,3,7,8}, Michael P. Morrissey^{4*}, William R. Sellers^{4*}, Robert Schlegel^{4*} & Levi A. Garraway^{1,2,3*}

The systematic translation of cancer genomic data into knowledge of tumour biology and therapeutic possibilities remains challenging. Such efforts should be greatly aided by robust preclinical model systems that reflect the genomic diversity of human cancers and for which detailed genetic and pharmacological annotation is available¹. Here we describe the Cancer Cell Line Encyclopedia (CCLE): a compilation of gene expression, chromosomal copy number and massively parallel sequencing data from 947 human cancer cell lines. When coupled with pharmacological profiles for 24 anticancer drugs across 479 of the cell lines, this collection allowed identification of genetic, lineage, and gene-expression-based predictors of drug sensitivity. In addition to known predictors, we found that plasma cell lineage correlated with sensitivity to IGF1 receptor inhibitors; AHR expression was associated with MEK inhibitor efficacy in *NRAS*-mutant lines; and *SLFN11* expression predicted sensitivity to topoisomerase inhibitors. Together, our results indicate that large, annotated cell-line collections may help to enable preclinical stratification schemata for anticancer agents. The generation of genetic predictions of drug response in the preclinical setting and their incorporation into cancer clinical trial design could speed the emergence of ‘personalized’ therapeutic regimens².

Human cancer cell lines represent a mainstay of tumour biology and drug discovery through facile experimental manipulation, global and detailed mechanistic studies, and various high-throughput applications. Numerous studies have used cell-line panels annotated with both genetic and pharmacological data, either within a tumour lineage^{3–5} or across multiple cancer types^{6–12}. Although affirming the promise of systematic cell line studies, many previous efforts were limited in their depth of genetic characterization and pharmacological interrogation.

To address these challenges, we generated a large-scale genomic data set for 947 human cancer cell lines, together with pharmacological profiling of 24 compounds across ~500 of these lines. The resulting collection, which we termed the Cancer Cell Line Encyclopedia (CCLE), encompasses 36 tumour types (Fig. 1a and Supplementary Table 1; see also <http://www.broadinstitute.org/ccle>). All cell lines were characterized by several genomic technology platforms. The mutational status of >1,600 genes was determined by targeted massively parallel sequencing, followed by removal of variants likely to be germline events (Supplementary Methods). Moreover, 392 recurrent mutations affecting 33

known cancer genes were assessed by mass spectrometric genotyping¹³ (Supplementary Table 2 and Supplementary Fig. 1). DNA copy number was measured using high-density single nucleotide polymorphism arrays (Affymetrix SNP 6.0; Supplementary Methods). Finally, messenger RNA expression levels were obtained for each of the lines using Affymetrix U133 plus 2.0 arrays. These data were also used to confirm cell line identities (Supplementary Methods and Supplementary Figs 2–4).

We next measured the genomic similarities by lineage between CCLE lines and primary tumours from Tumorscape¹⁴, expO, MILE and COSMIC data sets (Fig. 1b–d and Supplementary Methods). For most lineages, a strong positive correlation was observed in both chromosomal copy number and gene expression patterns (median correlation coefficients of 0.77, range = 0.52–0.94, $P < 10^{-15}$, for copy number, and 0.60, range = 0.29–0.77, $P < 10^{-15}$, for expression, respectively; Fig. 1b, c and Supplementary Tables 3 and 4), as has been described previously^{3–5,15}. A positive correlation was also observed for point mutation frequencies (median correlation coefficient = 0.71, range = –0.06–0.97, $P < 10^{-2}$ for all but 3 lineages; Supplementary Fig. 5), even when *TP53* was removed from the data set (median correlation coefficient = 0.64, range = –0.31–0.97, $P < 10^{-2}$ for all but 3 lineages; Fig. 1d and Supplementary Table 5). Thus, with relatively few exceptions (Supplementary Information), the CCLE may provide representative genetic proxies for primary tumours in many cancer types.

Given the pressing clinical need for robust molecular correlates of anticancer drug response, we incorporated a systematic framework to ascertain molecular correlates of pharmacological sensitivity *in vitro*. First, 8-point dose–response curves for 24 compounds (targeted and cytotoxic agents) across 479 cell lines were generated (Supplementary Tables 1 and 6, and Supplementary Methods). These curves were represented by a logistical sigmoidal function with a maximal effect level (A_{\max}), the concentration at half-maximal activity of the compound (EC_{50}), a Hill coefficient representing the sigmoidal transition, and the concentration at which the drug response reached an absolute inhibition of 50% (IC_{50}).

Broadly active compounds, exemplified by the HDAC inhibitor LBH589 (panobinostat), showed a roughly even distribution of A_{\max} and EC_{50} values across most cell lines (Fig. 2a). In contrast, the RAF inhibitor PLX4720 had a more selective profile: A_{\max} or EC_{50} values for most cell lines could be categorized as ‘sensitive’ or ‘insensitive’ to

¹The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. ²Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. ³Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁴Novartis Institutes for Biomedical Research, Cambridge, Massachusetts 02139, USA. ⁵Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA. ⁶Novartis Institutes for Biomedical Research, Emeryville, California 94608, USA. ⁷Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁸Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. [†]Present addresses: Novartis Institutes for Biomedical Research, Cambridge, Massachusetts 02139, USA (J.B.); Sage Bionetworks, 1100 Fairview Ave. N., Seattle, Washington 98109, USA (A.A.M.); Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA (M.F.B.).

*These authors contributed equally to this work.

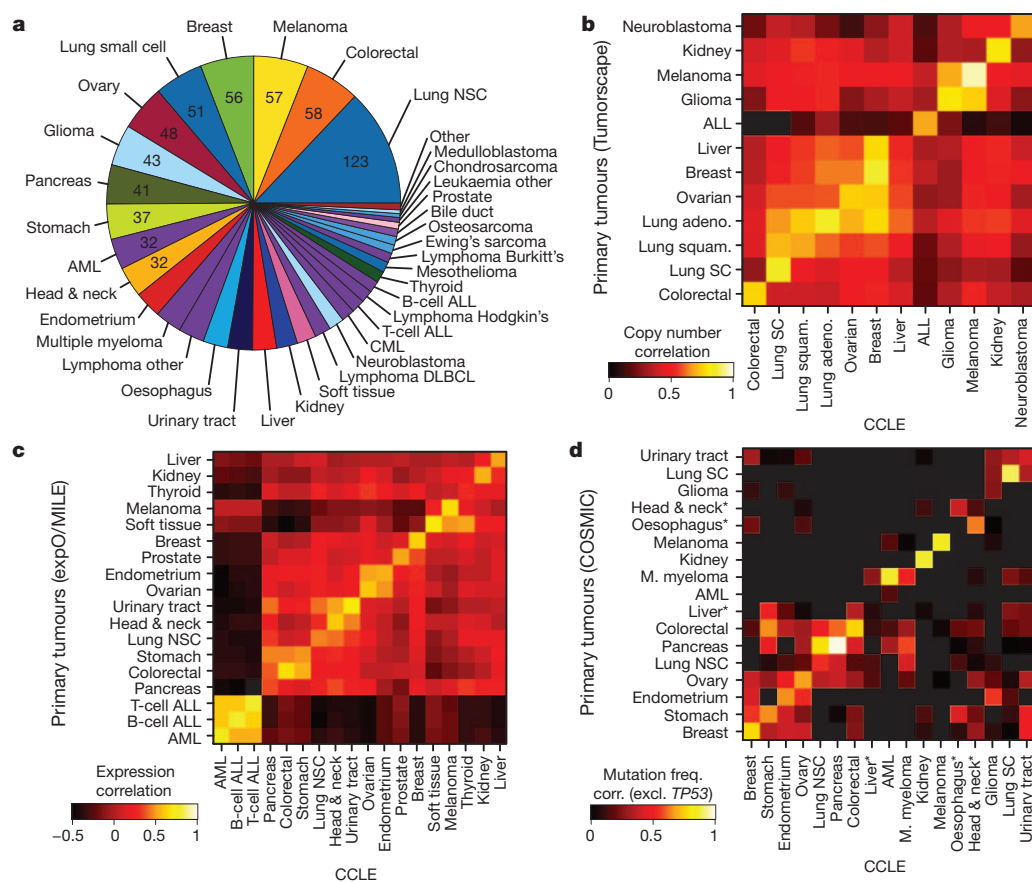


Figure 1 | The Cancer Cell Line Encyclopedia. **a**, Distribution of cancer types in the CCLE by lineage. **b**, Comparison of DNA copy-number profiles (GISTIC G-scores) between cell lines and primary tumours. The diagonal of the heat map shows the Pearson correlation between corresponding tumour types. Because cell lines and tumours are separate data sets, the correlation matrix is asymmetric: the top left showing how well the tumour features correlate with the average of the cell lines in a lineage, and the bottom right showing the converse. **c**, Comparison of mRNA expression profiles between cell lines and primary tumours. For each tumour type, the log fold change of the 5,000 most variable genes is calculated between that tumour type and all others. Pearson correlations between tumour type fold changes from primary tumours and cell lines are shown as a heat map. **d**, Comparison of point mutation frequencies between cell lines and primary tumours in COSMIC (v56), restricted to genes that are well represented in both sample sets but excluding *TP53*, which is highly prevalent in most tumour types. Pairwise Pearson correlations are shown as a heat map. Asterisk indicates that the correlations of oesophageal, liver, and head and neck cancer mutation frequencies are restored when including *TP53*.

PLX4720, with sensitive lines enriched for the *BRAF*^{V600E} mutation (Fig. 2a). To capture simultaneously the efficacy and potency of a drug, we designated an 'activity area' (Fig. 2b and Supplementary Fig. 6). The 24 compounds profiled showed wide variations in activity area, and those with similar mechanisms of action clustered together (Supplementary Fig. 7).

Genomic correlates of drug sensitivity may be extracted by predictive models using machine learning techniques^{6,10}. We therefore assembled all CCLE genomic data types into a matrix wherein each feature was converted to a z-score across all lines (Supplementary Methods). Next, we adapted a categorical modelling approach that used a naive Bayes classification and discrete sensitivity calls, or an elastic net regression analysis¹⁶ for continuous sensitivity measurements. Both approaches were applied to all compounds and genomic data with or without gene expression features (Supplementary Methods). Prediction performance was determined using tenfold cross-validation, and the elastic net features were bootstrapped to retain only those that were consistent across runs (Supplementary Methods).

Out of >50,000 input features, the regression-based analysis identified multiple known features as top predictors of sensitivity to several agents (Supplementary Table 7 and Supplementary Figs 8 and 9), with robust cross-validated performance (Supplementary Fig. 10 and 11). For example, activating mutations in *BRAF* and *NRAS* were among the top four predictors of sensitivity in models generated for the MEK inhibitor PD-0325901 (ref. 10) (Fig. 2c). Additional predictive features for MEK inhibition included expression of *PTEN*, *PTPN5* and *SPRY2* (which encodes a regulator of MAPK output). *KRAS* mutations were also identified, albeit with a lower predictive value (Fig. 2c, Supplementary Tables 8 and 9 and Supplementary Fig. 8).

Other top predictors included *EGFR* mutations and *ERBB2* amplification/overexpression for erlotinib⁸ and lapatinib¹⁷, respectively;

BRAF^{V600E} for RAF inhibitors (PLX4720 (ref. 18) and RAF265); *HGF* expression and *MET* amplification for the MET/ALK inhibitor PF-2341066 (ref. 19); and *MDM2* overexpression for Nutlin-3 (ref. 20) sensitivity. Variants affecting the *EXT2* gene, which encodes a glycosyltransferase involved in heparin sulphate biosynthesis, were significantly correlated with erlotinib effects (Supplementary Fig. 12). This observation is intriguing in light of a report linking heparin sulphate with erlotinib sensitivity²¹. In addition, *NQO1* expression was identified as the top predictive feature for sensitivity to the Hsp90 inhibitor 17-AAG, a quinone moiety metabolized by NAD(P)H:quinone oxidoreductase (*NQO1*). *NQO1* produces a high-potency intermediate (17-AAGH2)²², and has previously been identified as a potential biomarker for Hsp90 inhibitors²³.

Because some genetic/molecular alterations occur commonly in specific tumour types, lineage may become a confounding factor in predictive analyses. Indeed, a classifier built using the entire cell-line data set performed suboptimally when applied exclusively to melanoma-derived cell lines (Fig. 2d), whereas a model built with only melanoma cell lines performed better (Fig. 2d). Predictive features in the melanoma-only model showed a strong overexpression of genes regulated by the transcription factors MITF and SOX10 (Supplementary Table 10), which may also help predict RAF inhibitor drug sensitivity in melanoma cell lines.

Nonetheless, lineage emerged as the predominant predictive feature for several compounds. For example, elastic net studies of the HDAC inhibitor panobinostat identified haematological lineages as predictors of sensitivity (Fig. 2e and Supplementary Fig. 9). Interestingly, most clinical responses to panobinostat and related compounds (for example, vorinostat and romidepsin) have been observed in haematological cancers. Similarly, most multiple myeloma cell lines (12 of 14 lines tested) exhibited enhanced sensitivity to the IGF1 receptor inhibitor

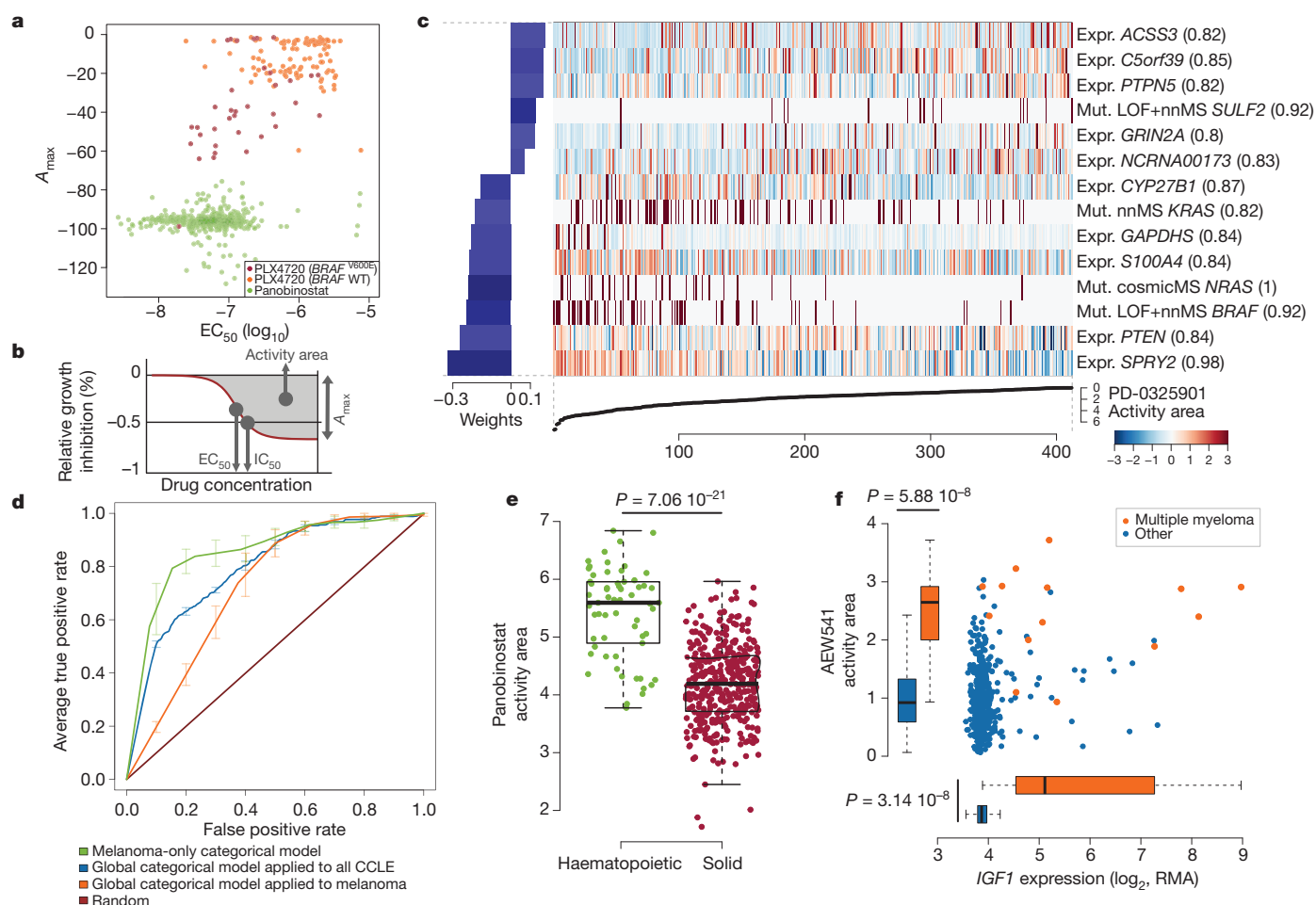


Figure 2 | Predictive modelling of pharmacological sensitivity using CCLE genomic data. **a, b,** Drug responses for panobinostat (green) and PLX4720 (orange/purple) represented by the high-concentration effect level (A_{\max}) and transitional concentration (EC_{50}) for a sigmoidal fit to the response curve (**b**). **c,** Elastic net regression modelling of genomic features that predict sensitivity to PD-0325901. The bottom curve indicates drug response, measured as the area over the dose-response curve (activity area), for each cell line. The central heat map shows the CCLE features in the model (continuous z-score for expression and copy number, dark red for discrete mutation calls), across all cell lines (x axis). Bar plot (left): weight of the top predictive features for sensitivity (bottom) or insensitivity (top). Parentheses indicate features present in >80% of models after bootstrapping. LOF, loss of function mutation; nnMS, non-neutral missense mutation (Supplementary Methods).

AEW541 (Fig. 2f and Supplementary Figs 8 and 9) and showed high *IGF1* expression (Fig. 2f). Interestingly, elevated *IGF1* expression also correlated with AEW541 sensitivity (Supplementary Fig. 9). The CCLE results indicate that multiple myeloma may be a promising indication for clinical trials of *IGF1* receptor inhibitors²⁴ and that these drugs may have enhanced efficacy in cancers with high *IGF1* or *IGF1R* expression.

Whereas *BRAF* and *NRAS* mutations are known single-gene predictors of sensitivity to MEK inhibitors, several 'sensitive' cell lines lacked mutations in these genes, whereas other lines harbouring these mutations were nonetheless 'insensitive' (Fig. 2c). The elastic net regression model derived from the subset of cell lines with validated *NRAS* mutations identified elevated expression of the *AHR* gene (which encodes the aryl hydrocarbon receptor) as strongly correlated with sensitivity to the MEK inhibitor PD-0325901 (Fig. 3a). This finding was interesting in light of previous studies indicating that a related MEK inhibitor (PD-98059) may also function as a direct *AHR* antagonist²⁵. We therefore hypothesized that the enhanced sensitivity of some *NRAS*-mutant cell lines to MEK inhibitors might relate to a coexistent dependence on *AHR* function.

d, Specificity and sensitivity (receiver operating characteristic curves) of cross-validated categorical models predicting the response to a MEK inhibitor, PD-0325901 (activity area). Mean true positive rate and standard deviation ($n = 5$) are shown when models are built using all lines (global categorical model, in blue and orange), or within only melanoma lines (green). **e,** Activity area values for panobinostat between cell lines derived from haematopoietic ($n = 61$) and solid tumours ($n = 387$). The middle bar, median; box, inter-quartile range; bars extend to $1.5\times$ the inter-quartile range. **f,** Distribution of activity area values for AEW541 relative to *IGF1* mRNA expression. Orange dots, multiple myeloma cell lines ($n = 14$); blue dots, cell lines from other tumour types ($n = 434$). Box-and-whisker plots show the activity area or mRNA expression distributions relative to each cell line type (line, median; box, inter-quartile range), with bars extending to $1.5\times$ the inter-quartile range.

To test this hypothesis, we first confirmed the correlation between *AHR* expression and sensitivity to MEK inhibitors in a subset of *NRAS*-mutant cell lines (Fig. 3b and Supplementary Fig. 13). Next, we performed short hairpin RNA (shRNA) knockdown of *AHR* in cell lines with high or low *AHR* expression (Fig. 3c). Silencing of *AHR* suppressed the growth of three *NRAS*-mutant cell lines with elevated *AHR* expression (Fig. 3d–f), but had no effect on the growth of two lines with low *AHR* expression (Fig. 3g, h). The growth inhibitory effect was confirmed with two additional shRNAs, where evidence for dose dependence was also apparent (Fig. 3i, j). We also tested the hypothesis that allosteric MEK inhibitors may suppress *AHR* function by measuring the effect of PD-0325901 and PD-98059 on endogenous *CYP1A1* mRNA, a transcriptional target of *AHR* in some contexts. Both compounds reduced *CYP1A1* levels in *NRAS*-mutant melanoma cells (IPC-298 and SK-MEL-2; Fig. 3k) but not in neuroblastoma cells (CHP-212; Fig. 3k), indicating that other factors may govern *CYP1A1* expression in the latter lineage. Together, these results suggest that *AHR* dependency may co-occur with MAP kinase activation in some *NRAS*-mutant cancer cells, and that elevated *AHR* may serve as a

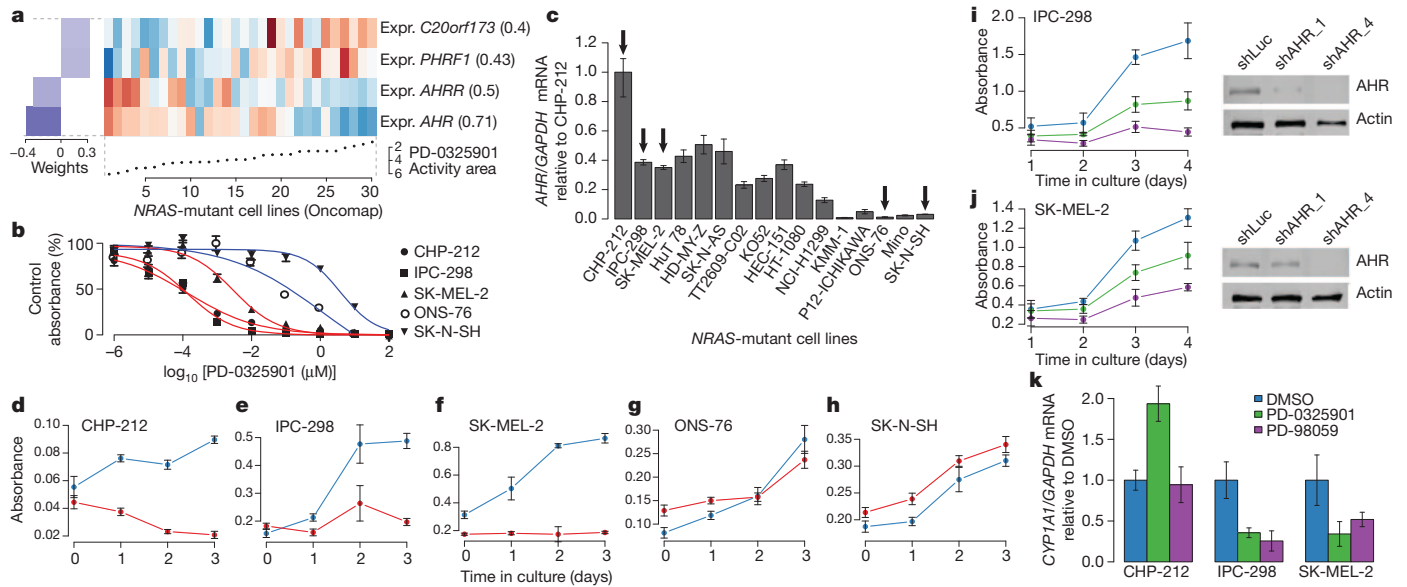


Figure 3 | AHR expression may denote a tumour dependency targeted by MEK inhibitors in NRAS-mutant cell lines. **a**, Predictive features for PD-0325901 sensitivity (using the 'varying baseline' activity area) in validated NRAS-mutant cell lines. **b**, Growth inhibition curves for NRAS-mutant cell lines expressing high (red) or low (blue) levels of AHR mRNA in the presence of the MEK inhibitor PD-0325901. **c**, Relative AHR mRNA expression across a panel of NRAS-mutant cell lines (arrows indicate cell lines where AHR dependency was analysed). **d–h**, Proliferation of NRAS-mutant cell lines displaying high (**d–f**) and low (**g, h**) AHR mRNA expression, after introduction of shRNAs against mechanistic biomarker for enhanced MEK inhibitor sensitivity in this setting.

We also looked for markers predictive of response to several conventional chemotherapeutic agents (Supplementary Fig. 7 and Supplementary Table 6) and identified *SLFN11* expression as the top correlate of sensitivity to irinotecan (Fig. 4a), a camptothecin analogue that inhibits the topoisomerase I (TOP1) enzyme. *SLFN11* expression

also emerged as the top predictor of topotecan sensitivity (another TOP1 inhibitor; Supplementary Figs 8 and 14). Overall, 12 of 16 lineages showed significant *SLFN11* associations for topotecan or irinotecan sensitivity (Pearson's $r \geq 0.2$, Supplementary Fig. 14b). This finding was independently validated using data from the NCI-60 collection (Supplementary Fig. 15). *SLFN11* knockdown did not affect steady-state growth sensitivity profiles (Supplementary Fig. 14d–f).

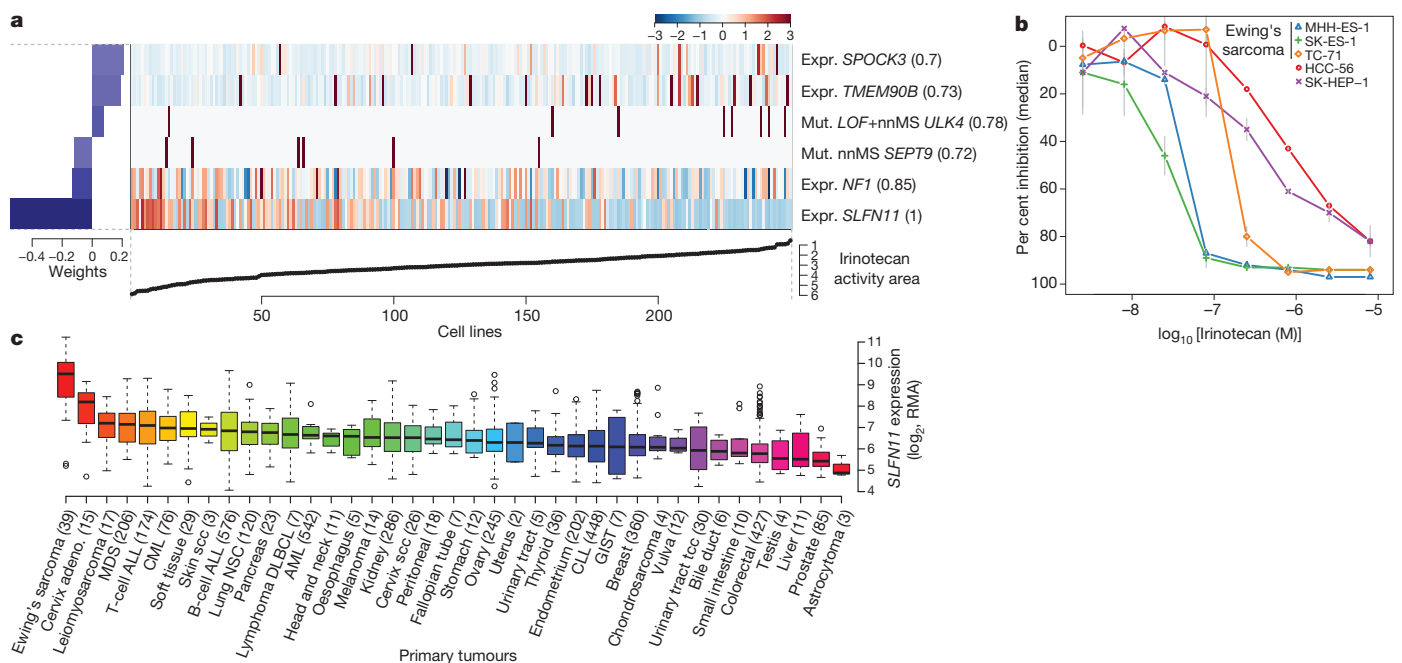


Figure 4 | Predicting sensitivity to topoisomerase I inhibitors. **a**, Elastic net regression analysis of genomic correlates of irinotecan sensitivity is shown for 250 cell lines. **b**, Dose-response curves for three Ewing's sarcoma cell lines (MSS-ES-1, SK-ES-1 and TC-71) and two control cell lines with low *SLFN11* expression (HCC-56 and SK-HEP-1). Grey vertical bars, standard deviation of the mean growth inhibition ($n = 2$). **c**, *SLFN11* expression across 4,103 primary tumours. Box-and-whisker plots show the distribution of mRNA expression for each subtype, ordered by the median *SLFN11* expression level (line), the inter-quartile range (box) and up to 1.5x the inter-quartile range (bars). Sample numbers (n) are indicated in parentheses.

the mean growth inhibition ($n = 2$). **c**, *SLFN11* expression across 4,103 primary tumours. Box-and-whisker plots show the distribution of mRNA expression for each subtype, ordered by the median *SLFN11* expression level (line), the inter-quartile range (box) and up to 1.5x the inter-quartile range (bars). Sample numbers (n) are indicated in parentheses.

All three Ewing's sarcoma cell lines screened showed both high *SLFN11* expression and sensitivity to irinotecan (Fig. 4b and Supplementary Fig. 14). Ewing's sarcomas also exhibited the highest *SLFN11* expression among 4,103 primary tumour samples spanning 39 lineages (Fig. 4c), suggesting that TOP1 inhibitors might offer an effective treatment option for this cancer type. Towards this end, several ongoing trials in Ewing's sarcoma are examining irinotecan-based combinations, or the addition of topotecan to standard regimens²⁶. For some lineages with high *SLFN11* expression (for example, cervical adenocarcinoma), topoisomerase inhibitors already comprise a standard chemotherapy regimen. In other tumours where topoisomerase inhibitors are commonly used (for example, colorectal and ovarian cancers), a range of *SLFN11* expression was observed, raising the possibility that high *SLFN11* expression might enrich for tumours more likely to respond. If confirmed in correlative clinical studies, *SLFN11* expression may offer a means to stratify patients for topoisomerase inhibitor treatment.

By assembling the CCLE, we have expanded the process of detailed annotation of preclinical human cancer models (<http://www.broadinstitute.org/ccle>). Genomic predictors of drug sensitivity revealed both known and novel candidate biomarkers of response. Even within genetically defined sub-populations—or when agents were broadly active without clear genetic targets—elastic net modelling studies identified key predictors or mechanistic effectors of drug response. Additional efforts that increase the scale and provide complementary types of information (for example, whole-genome/transcriptome sequencing, epigenetic studies, metabolic profiling or proteomic/phosphoproteomic analysis) should enable additional insights. In the future, comprehensive and tractable cell-line systems provided through this and other efforts²⁷ may facilitate numerous advances in cancer biology and drug discovery.

METHODS SUMMARY

A total of 947 independent cancer cell lines were profiled at the genomic level (data available at <http://www.broadinstitute.org/ccle> and Gene Expression Omnibus (GEO) using accession number GSE36139) and compound sensitivity data were obtained for 479 lines (Supplementary Table 11). Mutation information was obtained both by using massively parallel sequencing of >1,600 genes (Supplementary Table 12) and by mass spectrometric genotyping (OncoMap), which interrogated 492 mutations in 33 known oncogenes and tumour suppressors. Genotyping/copy number analysis was performed using Affymetrix Genome-Wide Human SNP Array 6.0 and expression analysis using the GeneChip Human Genome U133 Plus 2.0 Array. Eight-point dose–response curves were generated for 24 anticancer drugs using an automated compound-screening platform. Compound sensitivity data were used for two types of predictive models that used the naive Bayes classifier or the elastic net regression algorithm. The effects of *AHR* expression silencing on cell viability were assessed by stable expression of shRNA lentiviral vectors targeting either this gene or luciferase as control. The effect of compound treatment on *AHR* target gene expression was assessed by quantitative RT–PCR. A full description of the Methods is included in Supplementary Information.

Received 25 July 2011; accepted 1 March 2012.

- Caponigro, G. & Sellers, W. R. Advances in the preclinical testing of cancer therapeutic hypotheses. *Nature Rev. Drug Discov.* **10**, 179–187 (2011).
- MacConaill, L. E. & Garraway, L. A. Clinical implications of the cancer genome. *J. Clin. Oncol.* **28**, 5219–5228 (2010).
- Lin, W. M. *et al.* Modeling genomic diversity and tumor dependency in malignant melanoma. *Cancer Res.* **68**, 664–673 (2008).
- Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).
- Sos, M. L. *et al.* Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions. *J. Clin. Invest.* **119**, 1727–1740 (2009).
- Dry, J. R. *et al.* Transcriptional pathway signatures predict MEK addiction and response to selumetinib (AZD6244). *Cancer Res.* **70**, 2264–2273 (2010).
- Garraway, L. A. *et al.* Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122 (2005).
- Greshock, J. *et al.* Molecular target class is predictive of *in vitro* response profile. *Cancer Res.* **70**, 3677–3686 (2010).
- McDermott, U. *et al.* Identification of genotype-correlated sensitivity to selective kinase inhibitors by using high-throughput tumor cell line profiling. *Proc. Natl Acad. Sci. USA* **104**, 19936–19941 (2007).

- Solit, D. B. *et al.* BRAF mutation predicts sensitivity to MEK inhibition. *Nature* **439**, 358–362 (2006).
- Staunton, J. E. *et al.* Chemosensitivity prediction by transcriptional profiling. *Proc. Natl Acad. Sci. USA* **98**, 10787–10792 (2001).
- Weinstein, J. N. *et al.* An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343–349 (1997).
- Thomas, R. K. *et al.* High-throughput oncogene mutation profiling in human cancer. *Nature Genet.* **39**, 347–351 (2007).
- Beroukhi, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).
- Ross, D. T. *et al.* Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genet.* **24**, 227–235 (2000).
- Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* **67**, 301–320 (2005).
- Konecny, G. E. *et al.* Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res.* **66**, 1630–1639 (2006).
- Tsai, J. *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl Acad. Sci. USA* **105**, 3041–3046 (2008).
- Zou, H. Y. *et al.* An orally available small-molecule inhibitor of c-Met, PF-2341066, exhibits cytoreductive antitumor efficacy through antiproliferative and antiangiogenic mechanisms. *Cancer Res.* **67**, 4408–4417 (2007).
- Müller, C. R. *et al.* Potential for treatment of liposarcomas with the MDM2 antagonist Nutlin-3A. *Int. J. Cancer* **121**, 199–205 (2007).
- Nishio, M. *et al.* Serum heparan sulfate concentration is correlated with the failure of epidermal growth factor receptor tyrosine kinase inhibitor treatment in patients with lung adenocarcinoma. *J. Thorac. Oncol.* **6**, 1889–1894 (2011).
- Guo, W. *et al.* Formation of 17-allylamino-demethoxygeldanamycin (17-AAG) hydroquinone by NAD(P)H:quinone oxidoreductase 1: role of 17-AAG hydroquinone in heat shock protein 90 inhibition. *Cancer Res.* **65**, 10006–10015 (2005).
- Kelland, L. R., Sharp, S. Y., Rogers, P. M., Myers, T. G. & Workman, P. DT-Diaphorase expression and tumor cell sensitivity to 17-allylamino, 17-demethoxygeldanamycin, an inhibitor of heat shock protein 90. *J. Natl Cancer Inst.* **91**, 1940–1949 (1999).
- Moreau, P. *et al.* Phase I study of the anti insulin-like growth factor 1 receptor (IGF-1R) monoclonal antibody, AVE1642, as single agent and in combination with bortezomib in patients with relapsed multiple myeloma. *Leukemia* **25**, 872–874 (2011).
- Reiners, J. J. Jr, Lee, J. Y., Clift, R. E., Dudley, D. T. & Myrand, S. P. PD98059 is an equipotent antagonist of the aryl hydrocarbon receptor and inhibitor of mitogen-activated protein kinase kinase. *Mol. Pharmacol.* **53**, 438–445 (1998).
- Wagner, L. M. *et al.* Temozolomide and intravenous irinotecan for treatment of advanced Ewing sarcoma. *Pediatr. Blood Cancer* **48**, 132–139 (2007).
- Garnett, M. J. *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* <http://dx.doi.org/10.1038/nature11005> (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff of the Biological Samples Platform, the Genetic Analysis Platform and the Sequencing Platform at the Broad Institute. We thank S. Banerji, J. Che, C. M. Johannessen, A. Su and N. Wagle for advice and discussion. We are grateful for the technical assistance and support of G. Bonamy, R. Bruschi III, E. Gelfand, K. Gravelin, T. Huynh, S. Kehoe, K. Matthews, J. Nedzel, L. Niu, R. Pinchback, D. Roby, J. Slind, T. R. Smith, L. Tan, V. Trinh, C. Vickers, G. Yang, Y. Yao and X. Zhang. The Cancer Cell Line Encyclopedia project was enabled by a grant from the Novartis Institutes for Biomedical Research. Additional funding support was provided by the National Cancer Institute (M.M., L.A.G.), the Starr Cancer Consortium (M.F.B., L.A.G.), and the NIH Director's New Innovator Award (L.A.G.).

Author Contributions For the work described herein, J.B. and G.C. were the lead research scientists; N.S., K.V. and A.M.M. were the lead computational biologists; M.P.M., W.R.S., R.S. and L.A.G. were the senior authors. J.B., G.C., S.K., P.M., J.M., J.T., A.S., N.L. and K.A. performed cell-line procurement and processing; P.M. and K.A. performed or directed nucleic acid extraction and quality control; S.G., W.W. and S.B.G. performed or directed genomic data generation; C.J.W., F.A.M., E.B.-F., I.H.E., P.A., M.D.S., K.J. and V.E.M. performed pharmacological data generation; N.S., K.V., G.V.K., A.R., M.F.B., J.C., G.K.Y., M.D.J., T.L., M.R. and G.G. contributed to software development; N.S., K.V., A.A.M., J.L., G.V.K., D.S., A.R., M.L., M.F.B., A.K., P.R., J.C., G.K.Y., J.Y., M.D.J., L.W., C.H., E.P., J.P.M., V.C. and M.P.M. performed computational biology and bioinformatics analysis; J.B., G.C., N.S., L.M., J.E.M., J.J.-V., M.P.M., W.R.S., R.S. and L.A.G. performed biological analysis and interpretation; N.S., K.V., A.A.M., J.L., A.R., M.L., L.M., A.K., J.J.-V., J.C., G.K.Y. and J.Y. prepared figures and tables for the main text and Supplementary Information; J.B., G.C., N.S., K.V., A.A.M., J.L., G.V.K., J.J.-V., M.P.M. and L.A.G. wrote and edited the main text and Supplementary Information; J.B., G.C., N.S., K.V., S.K., C.J.W., J.L., S.M., C.S., R.C.O., T.L., L.M.C., W.W., M.R., N.L., S.B.G., K.A. and V.C. performed project management; J.P.M., V.E.M., B.L.W., J.P., M.W., P.F., J.L.H., M.M. and T.R.G. contributed project oversight and advisory roles; and M.P.M., W.R.S., R.S. and L.A.G. provided overall project leadership.

Author Information Data have been deposited in the Gene Expression Omnibus (GEO) using accession number GSE36139 and are also available at <http://www.broadinstitute.org/ccle>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.A.G. (Levi_Garraway@dfci.harvard.edu) or R.S. (robert.schlegel@novartis.com).

Deregulated MYC expression induces dependence upon AMPK-related kinase 5

Lidan Liu^{1*}, Jannes Ulbrich^{1*}, Judith Müller^{1†}, Torsten Wüstefeld^{2,3}, Lukas Aeberhard^{4†}, Theresia R. Kress^{1†}, Nathiya Muthalagu¹, Lukas Rycak⁵, Ramona Rudalska², Roland Moll⁶, Stefan Kempa⁴, Lars Zender^{2,3}, Martin Eilers¹ & Daniel J. Murphy¹

Deregulated expression of the MYC oncoprotein contributes to the genesis of many human tumours, yet strategies to exploit this for a rational tumour therapy are scarce. MYC promotes cell growth and proliferation, and alters cellular metabolism to enhance the provision of precursors for phospholipids and cellular macromolecules^{1,2}. Here we show in human and murine cell lines that oncogenic levels of MYC establish a dependence on AMPK-related kinase 5 (ARK5; also known as NIAK1) for maintaining metabolic homeostasis and for cell survival. ARK5 is an upstream regulator of AMPK and limits protein synthesis via inhibition of the mammalian target of rapamycin 1 (mTORC1) signalling pathway. ARK5 also maintains expression of mitochondrial respiratory chain complexes and respiratory capacity, which is required for efficient glutamine metabolism. Inhibition of ARK5 leads to a collapse of cellular ATP levels in cells expressing deregulated MYC, inducing multiple pro-apoptotic responses as a secondary consequence. Depletion of ARK5 prolongs survival in MYC-driven mouse models of hepatocellular carcinoma, demonstrating that targeting cellular energy homeostasis is a valid therapeutic strategy to eliminate tumour cells that express deregulated MYC.

To identify kinases that are specifically required for the viability of cells expressing deregulated MYC, we used U2OS cells expressing c-MYC fused to the oestrogen receptor ligand binding domain (MYC-ER) (Fig. 1a). Activation of MYC-ER by 4-hydroxytamoxifen (OHT) had little effect on apoptosis when cells were grown at low density in the presence of growth factors. Under these conditions, we performed a short interfering (si)RNA screen of the human kinome, using automated microscopy to identify siRNAs that induced poly-ADP-ribose-polymerase cleavage specifically in the presence of OHT. This screen yielded two hits, ARK5 and AMPK (Supplementary Table 1).

Depletion of ARK5 induced the accumulation of MYC-expressing cells that stained positive for annexin V and propidium iodide (Fig. 1a and Supplementary Fig. 1a). Similarly, expressing different short hairpin (sh)RNAs targeting ARK5 induced levels of MYC-dependent death that correlated with the degree of knockdown (Fig. 1b). Titration of OHT revealed that levels of MYC that cause a dependence on ARK5 are higher than those required to promote proliferation (Supplementary Fig. 1b). Depletion of ARK5 induced death in U2OS cells constitutively expressing MYC and suppressed propagation of MRC5 fibroblasts in a MYC-dependent manner (Fig. 1c and Supplementary Fig. 1c). Expression of murine ARK5, which is not targeted by the shRNAs used, prevented death upon depletion of human ARK5 (Fig. 1d). This rescue required LKB1-dependent phosphorylation of T212, but not AKT-dependent phosphorylation of S601 (refs 3, 4). Mutation of K85 within the ATP-binding domain blocked the ability of murine ARK5 to prevent death, demonstrating that rescue requires ARK5 catalytic activity. Accordingly,

a small-molecule inhibitor of ARK5, BX795, mimicked the effects of ARK5 depletion (Fig. 1e and Supplementary Fig. 1d–f).

Ectopic expression of BCL2 or MCL1, which protect cells from apoptosis induced by growth-factor deprivation, failed to alleviate the dependence on ARK5 (Supplementary Fig. 2a and data not shown). Depletion of ARK5 did not enhance the pro-apoptotic activity of E2F1 or E2F2 (Supplementary Figs 2b, c). ARK5 regulates the Hippo pathway, as it destabilizes LATS1 (Supplementary Figs 1d and 2d). However, co-depletion of LATS1 had no effect on death upon depletion of ARK5 (Supplementary Fig. 2d). Furthermore, depletion of ARK5 had little effect on MYC-dependent target gene activation (Supplementary Fig. 2e). To identify relevant effector pathways of ARK5, we characterized the effects of ARK5 depletion on cell physiology. Depletion of ARK5 delayed progression through all phases of the cell cycle (Fig. 2a, b). ARK5-depleted cells were larger than controls during S and G2 phase, demonstrating that ARK5 restricts cell growth (Fig. 2c and Supplementary Fig. 2f).

AMPK inhibits the mTORC1 pathway that controls anabolic cell growth⁸. Different shRNAs targeting AMPK induced death in a MYC-dependent manner (Supplementary Fig. 3a). MYC promotes multiple anabolic processes and might thereby strain energy supplies and activate AMPK. Indeed, we observed a progressive increase in T172-phosphorylated AMPK⁹ and a progressive decrease in mTORC1 activity in MYC-expressing cells (Fig. 2d, e). Depletion of ARK5 ablated activation of AMPK in response to MYC and to the AMP analogue aminoimidazole carboxamide ribonucleotide (AICAR) (Fig. 2f and Supplementary Fig. 3b). Consistently, depletion of ARK5 enhanced overall protein synthesis in a manner that was sensitive to the mTORC1 inhibitor rapamycin (Fig. 2g). Proteomic analysis demonstrated that ARK5 is required for protecting the β 1 subunit of AMPK, which activates AMPK in response to alterations in the ATP:AMP ratio, from proteasomal degradation (Fig. 2f, Supplementary Fig. 3c and Supplementary Table 2)¹⁰. These findings suggested that unrestrained mTORC1 activity contributes to the death of MYC-expressing cells when depleted of ARK5 (ref. 11). Accordingly, addition of rapamycin protected cell viability in the absence of ARK5 or AMPK (Fig. 3a and Supplementary Fig. 3d). Similarly, addition of OSI-027, an inhibitor of mTORC1 and mTORC2, prevented death of ARK5-depleted cells, excluding the possibility that AKT signalling accounts for the blockade of death (Supplementary Fig. 3e)^{12,13}.

Depletion of ARK5 in MYC-expressing cells led to a progressive rapamycin-sensitive decline in ATP levels, which preceded death (Fig. 3a, b and Supplementary Fig. 4). Correlating with increased protein synthesis, depletion of ARK5 enhanced uptake of glutamine (Supplementary Fig. 5a). Because MYC-transformed cells depend on glutamine^{14–16}, loss of ATP might be secondary to depletion of glutamine in the medium. However, daily addition of either glutamine

¹Theodor Boveri Institute, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany. ²Helmholtz Centre for Infection Research, Inhoffenstrasse 7, 38124 Braunschweig, Germany.

³Department of Gastroenterology, Hepatology and Endocrinology, Medical School Hannover, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany. ⁴Max-Delbrück-Center for Molecular Medicine, Berlin Institute for Medical Systems Biology, Robert Rössle Strasse 10, Berlin 13125, Germany. ⁵Institute of Molecular Biology and Tumor Research, Emil-Mannkopff-Str. 2, 35037 Marburg, Germany. ⁶Institute for Pathology, University of Marburg, Baldingerstrasse, 35033 Marburg, Germany. [†]Present addresses: Netherlands Cancer Institute (NKI), Department of Molecular Genetics, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands (J.M.); Robert Koch Institute, Nordufer 20, 13353 Berlin, Germany (L.A.); European Institute of Oncology, IFOM-IEO Campus, Via Adamello, 16 - 20139 Milan, Italy (T.R.K.).

*These authors contributed equally to this work.

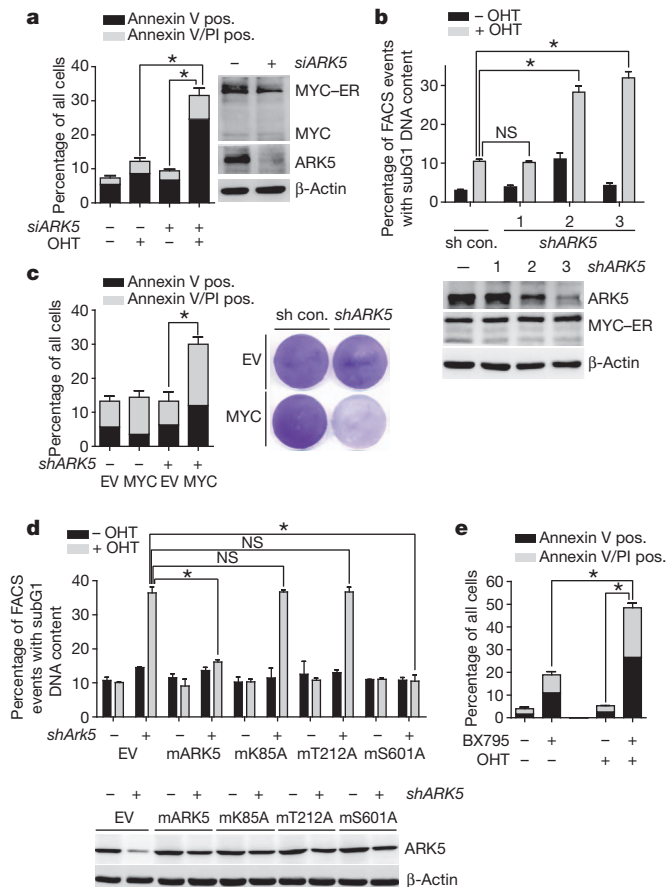


Figure 1 | Synthetic lethality of MYC deregulation and inhibition of ARK5 inhibition. **a**, Depletion of ARK5 induces MYC-dependent death. U2OS MYC-ER cells were transfected with control siRNA or *siARK5* and treated with OHT or solvent. The graph shows the percentage of cells staining positive (pos.) for annexin V only (black) or for both annexin V and propidium iodide (grey). Results were consistent in three independent experiments. Immunoblots document expression of ARK5, MYC-ER and endogenous MYC. Results are shown as mean plus standard deviation (s.d.) of biological triplicates from one representative experiment (here: for cumulative death, annexin V positive plus annexin V and propidium iodide (PI) double positive), except where expressly stated. **P* values < 0.01; NS, not statistically significant. **b**, shRNA depletion of ARK5 induces MYC-dependent death. Top, percentage of U2OS MYC-ER cells with subG1 DNA content after expression of control shRNA (sh con.) or *shARK5*. Results are averaged (\pm s.d.) from three independent experiments. Bottom, immunoblots documenting expression of ARK5, MYC-ER and β -actin. **c**, ARK5 depletion kills cells expressing constitutive MYC. FACS analysis (left) or crystal violet staining (right) of control and MYC-expressing U2OS cells, 4 days after retroviral expression of *shARK5* (sequence 3 shown in panel b). EV, empty vector. Results are representative of two independent experiments. **d**, ARK5 kinase activity is required to prevent MYC-dependent death. U2OS MYC-ER cells expressing the indicated murine ARK5 point mutant proteins, superinfected with retroviruses expressing *shArk5*. Results are representative of three independent experiments. The immunoblot documents shRNA-resistant expression of murine ARK5. **e**, Pharmacological inhibition of ARK5 drives MYC-dependent death. U2OS MYC-ER cells were treated once with 50 nM BX795 or solvent, in the presence or absence of OHT and analysed after 5 days. Results are representative of four independent experiments.

or glucose failed to prevent death (Supplementary Fig. 5b). Furthermore, low levels of autophagy were observed in U2OS MYC-ER cells upon activation of MYC, and depletion of ARK5 moderately enhanced the accumulation of LC3 punctae, indicative of autophagy, suggesting that altered autophagy does not account for the loss of ATP (not shown)¹⁷. Finally, activation of MYC had only small effects on protein (Fig. 2g) and DNA synthesis (Supplementary Fig. 1b) that are unlikely to alone account for the ATP loss in MYC-expressing cells.

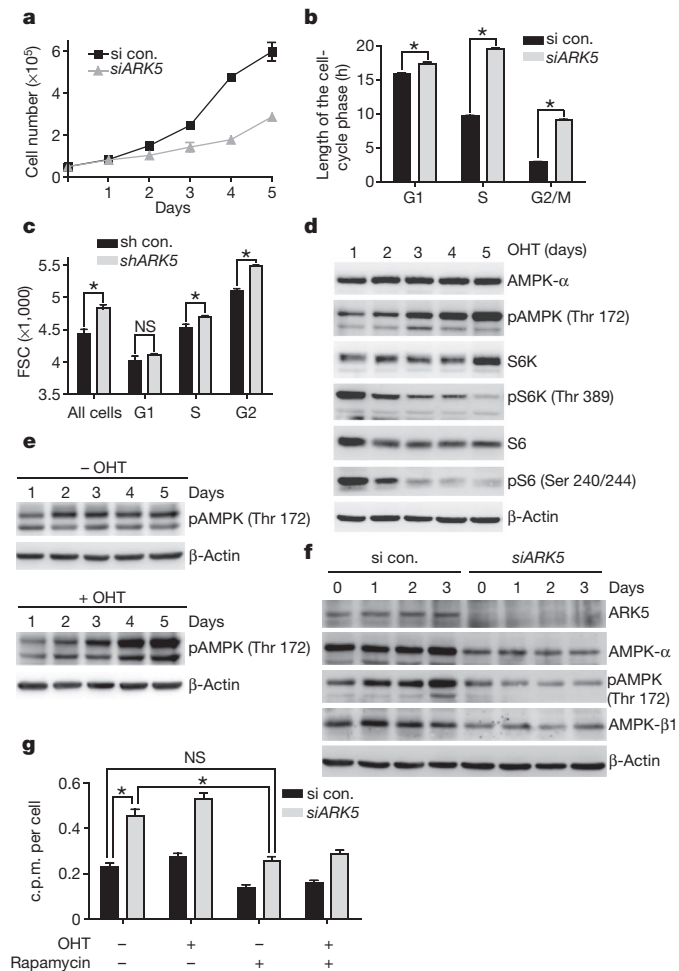


Figure 2 | ARK5 restrains cell growth and the mTOR pathway. **a**, Depletion of ARK5 retards cell proliferation. U2OS cells were transfected with *siARK5* and counted daily. Results are representative of four independent experiments. si con., control siRNA. **b**, Length of cell-cycle phases in ARK5-depleted U2OS cells. The percentage of cells in each phase of the cell cycle was determined by BrdU fluorescence-activated cell sorting (FACS). Using the doubling time from the experiment described in **a**, the length of each cell-cycle phase was calculated. Consistent results were obtained in two independent experiments. **c**, ARK5-depleted U2OS MYC-ER cells are larger than controls. ARK5-depleted and control cells were stained with Vybrant Dye violet, gated by DNA content and size was determined for G1, S and G2 cells. sh con., control shRNA. FSC, forward scatter. **d**, MYC-dependent activation of AMPK during culture of U2OS cells. Lysates were prepared from U2OS MYC-ER cells cultured in the presence of OHT and probed with the indicated antibodies. Results are representative of three independent experiments. p, phospho. **e**, Lysates from U2OS MYC-ER cells cultured with or without OHT were probed with anti-T172-phosphorylated AMPK- α 1 and β -actin. **f**, ARK5 is required for activation of AMPK in response to MYC. U2OS MYC-ER cells transfected with control or *siARK5*, treated as per panel **e**. Results are representative of three independent experiments. AMPK- β 1 denotes the β 1 subunit of AMPK. **g**, ARK5 restricts global protein synthesis. U2OS MYC-ER cells, expressing *shARK5* or control shRNA, treated with or without OHT and/or rapamycin and labelled after 24 h with ³H-leucine. Label incorporation per cell is shown. c.p.m., counts per minute. Consistent results were obtained using ³H-lysine and across three independent experiments in total. Results are shown as mean plus s.d.; **P* values < 0.01; NS, not statistically significant.

These results were recapitulated in a murine hepatocellular carcinoma line that expresses MYC and inducible shRNAs targeting *ARK5* (see later). We pulse-labelled these cells with ¹³C-glucose or ¹³C-glutamine and traced the carbon flow. Consistent with previous demonstrations that deregulated MYC diverts glucose away from mitochondrial metabolism, flow from glucose into the tricarboxylic acid (TCA) cycle

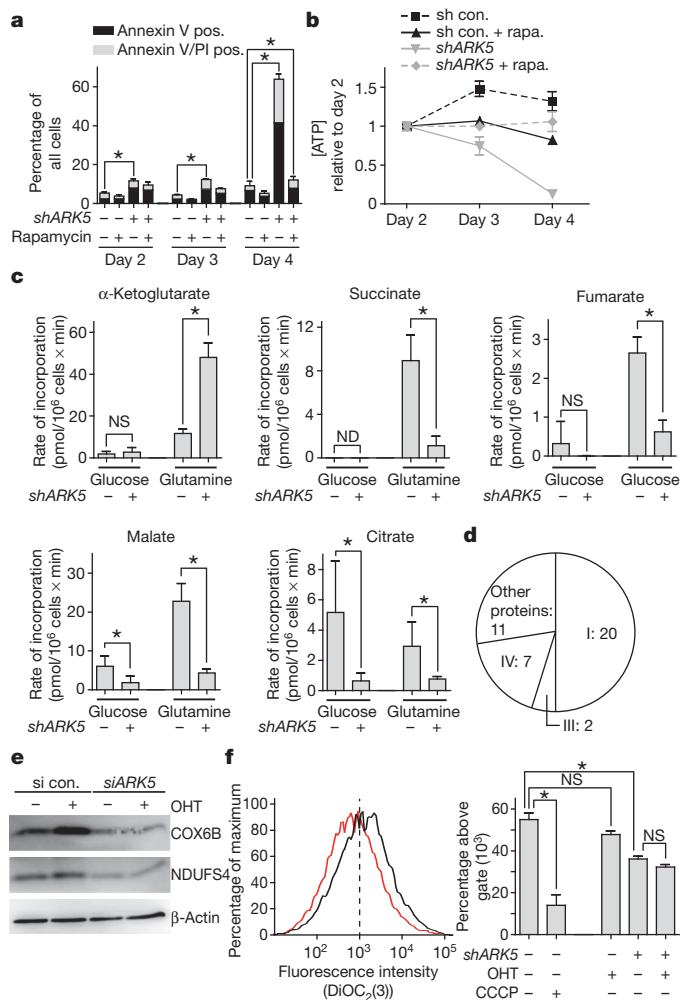


Figure 3 | Failure to restrain mTOR contributes to death in cells expressing deregulated MYC. **a**, Rapamycin protects MYC-overexpressing cells from ARK5 depletion. U2OS MYC-ER cells, expressing *shARK5* or control shRNA (sh con.) and cultured in the presence of OHT with or without 100 nM rapamycin, as indicated. The percentage of apoptotic cells was determined daily by FACS and consistent results obtained in two independent time-course analyses. **b**, Depletion of ARK5 induces a MYC-dependent collapse of cellular ATP levels. U2OS MYC-ER cells transfected with control siRNA or *siARK5* and treated with OHT. Viable cells were harvested daily. Levels of ATP are plotted relative to day 2. rapa., rapamycin. Results are representative of two experiments. **c**, Gas chromatography-mass spectrometry (GC-MS) analysis of carbon flow in MYC-transformed HCCs with or without ARK5. HCCs expressing MYC, AKT and doxycycline-inducible *shARK5* were labelled with ^{13}C -glucose or ^{13}C -glutamine 48 h after induction of *shARK5*, as indicated. Graphs show the rate of label incorporation into TCA cycle intermediates (picomoles/ 1×10^6 cells per minute). Mean results and s.d. from three biological replicates, each with two technical replicates, are shown. ND, not detected. **d**, Proteomic analysis of ARK5 depletion. The chart summarizes results of a stable isotope labelling by amino acids in cell culture (SILAC) experiment measuring the abundance of total cellular proteins before and after depletion of ARK5. Subunit constituents of the mitochondrial respiratory complexes I, III and IV account for 29 of the 40 most suppressed proteins in ARK5-depleted cells. **e**, Induction of respiratory chain components by MYC is overridden by ARK5 depletion. Immunoblots show U2OS MYC-ER cells expressing *shARK5* or control shRNA, treated with or without OHT for 24 h. Results are representative of two independent experiments. **f**, ARK5 depletion reduces mitochondrial membrane potential. U2OS MYC-ER cells expressing *shARK5* or control shRNA, treated for 48 h with OHT, or acutely with carbonyl cyanide m-chlorophenyl hydrazone (CCCP), labelled with 3,3'-diethyloxycarbocyanine iodide (DiOC₂(3)) and analysed by FACS. Left, representative FACS distribution of ARK5-depleted, OHT-treated cells (red) compared with untreated control cells (black). Right, percentage of cells with fluorescent intensity above 10^3 in each condition. Results are representative of two independent experiments. **P* values < 0.01; NS, not statistically significant.

was low and most glucose was converted to lactate (Fig. 3c and Supplementary Fig. 5c, d)^{14,16}. Depletion of ARK5 had little effect on glycolysis (Supplementary Fig. 5c). In line with enhanced uptake of glutamine, cells depleted of ARK5 exhibited increased flow from ^{13}C -glutamine to α -ketoglutarate (Fig. 3c and Supplementary Fig. 5d). Depletion of ARK5 inhibited the entry of α -ketoglutarate into the TCA cycle, as flow into succinate, fumarate and malate was suppressed (Fig. 3c). Addition of membrane-permeable di-methyl-2-oxoglutarate did not rescue ARK5-depleted cells (data not shown). A reduction in TCA cycling was also evident from reduced flow into citrate, irrespective of the carbon source (Fig. 3c).

Depletion of ARK5 had no effect on the expression of α -ketoglutarate dehydrogenase subunits at either protein or messenger RNA levels (Supplementary Fig. 6a). A proteomic analysis revealed downregulation of multiple subunits of complexes I, III and IV of the mitochondrial respiratory chain (Fig. 3d and Supplementary Table 3; Gene Ontology term Benjamin enrichment score, 2×10^{-48}). This regulation is post-transcriptional, specific and independent of the proteasome (Supplementary Tables 2, 4 and Supplementary Figs 3c, 6b). Immunoblot analyses confirmed previous observations that MYC upregulates expression of respiratory chain proteins and showed that depletion of ARK5 overrides this effect (Fig. 3e and Supplementary Fig. 6c)^{18,19}. Depletion of ARK5 also lowered the mitochondrial membrane potential (Fig. 3f). Consistently, depletion of ARK5 reduced oxygen consumption, indicating that an inability to oxidize NADH limits entry of α -ketoglutarate into the TCA cycle (Supplementary Fig. 6d). Therefore, ARK5 is required to maintain sufficient respiratory capacity to sustain glutamine consumption in MYC transformed cells.

To ascertain if the restriction of mitochondrial respiration contributes to death, we limited mitochondrial function by depriving cells of nutrients or by blocking oxidative phosphorylation (Supplementary Fig. 7a). Cells cultured in the absence of nutrients and in the presence of oligomycin had undetectable levels of ATP and underwent cell death regardless of MYC expression. In the absence of nutrients, ATP levels declined more rapidly in MYC-expressing cells and these specifically underwent rapid cell death (Supplementary Fig. 7a-c). Consistently, depletion of ARK5 in MYC-expressing cells induced multiple hallmarks of low cellular ATP levels or limited electron transport, including an unfolded protein response, accumulation of DNA damage and enhanced levels of reactive oxygen species (Supplementary Fig. 7d-g).

Available databases document increased levels of ARK5 mRNA in hepatocellular and pancreatic carcinoma (Supplementary Fig. 8a). Immunohistochemistry revealed elevated ARK5 expression in the majority of colon, pancreatic and hepatocellular carcinoma cases and a high overlap with MYC expression (Fig. 4a and Supplementary Fig. 8b, c). Accordingly, depletion of ARK5 suppressed proliferation of 5/14 human tumour cell lines (Supplementary Fig. 8d). Depletion of ARK5 or AMPK induced cell death in Ls174T colon carcinoma cells, which harbour a mutant β -catenin gene that drives MYC expression²⁰. Death was paralleled by loss of ATP and prevented by rapamycin (Supplementary Figs 9a, b). Co-depletion of MYC also prevented death, demonstrating that deregulated MYC establishes the dependence on ARK5 (Supplementary Fig. 9c). Moreover, depletion of ARK5 suppressed tumour formation by Ls174T cells (Supplementary Fig. 9d).

To test the therapeutic efficacy of ARK5 depletion, we transplanted murine *p53*^{-/-} hepatoma cells that express MYC and AKT together with two different doxycycline-inducible *Ark5* shRNAs under the liver capsule to induce orthotopic carcinomas (Supplementary Fig. 10a)^{21,22}. *In vitro*, depletion of ARK5 led to culture collapse, preceded by loss of ATP (Fig. 4b, c, Supplementary Fig. 10b and data not shown). Addition of rapamycin restored ATP levels and allowed sustained culture of ARK5-depleted cells (Supplementary Fig. 10c). Upon transplantation, hepatocellular carcinomas (HCCs) developed in 8/8 untreated control mice, but only 1/10 mice treated with doxycycline (two-sided

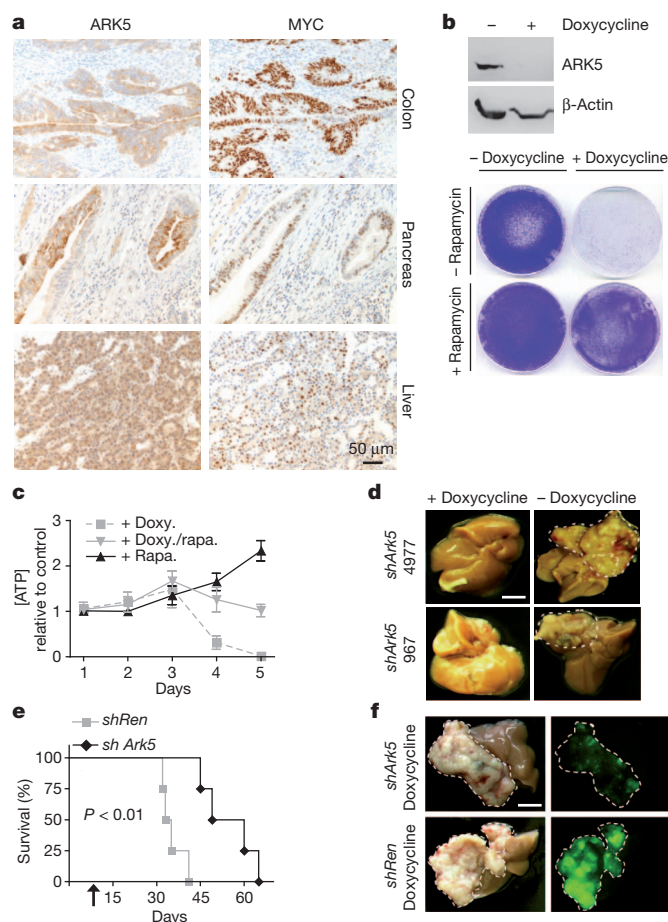


Figure 4 | Targeting ARK5 as a therapeutic strategy in hepatocellular carcinoma. **a**, ARK5 is expressed in human cancers. Samples from human colorectal carcinoma ($n = 26$), pancreatic ductal adenocarcinoma ($n = 21$) and hepatocellular carcinoma ($n = 11$) were analysed for expression of ARK5 and MYC by immunohistochemistry (scale bar, 50 μ m). **b**, Depletion of ARK5 suppresses propagation of HCC cells in a rapamycin-dependent manner. Top, immunoblots; bottom, crystal-violet-stained cultures of these cells. Consistent results were obtained in four independent experiments. **c**, ATP collapses after ARK5 depletion in HCC cells. ATP was measured daily in cells grown with or without doxycycline and/or rapamycin, as indicated. Results are plotted relative to untreated cultures. Consistent results were obtained in two independent experiments. Data are presented as mean \pm s.d. **d**, Suppression of ARK5 prevents tumorigenesis in a model of HCC. Images of representative livers from mice transplanted with HCC cells carrying either of two doxycycline-inducible *Ark5* shRNA vectors and treated with ($n = 5$ mice for each shRNA) or without ($n = 4$ mice for each shRNA) doxycycline from the time of transplantation. Tumours are outlined with a hashed border. Scale bar, 5 mm. **e**, Kaplan–Meier diagram documenting survival of mice orthotopically injected with 1×10^6 HCC cells. Expression of *shArk5* ($n = 4$ mice) or control shRNA ($n = 4$ mice) was induced 9 days after transplantation (arrow). Statistical comparison of Kaplan–Meier curves is based on the log-rank test. **f**, Selection against *shArk5* during HCC development. The panel shows representative examples of the liver of mice expressing *shArk5* or control shRNA (left) and GFP staining documenting selection against GFP-positive tumour cells expressing *shArk5* (right). Scale bar, 5 mm.

P value = 0.0004; Fig. 4d). Depletion of ARK5 also provided a survival advantage relative to cells expressing a shRNA targeting luciferase in an intervention study, in which tumours were allowed to develop before doxycycline was added ($P < 0.01$; Fig. 4e). The vector that drives the shRNA encodes GFP, reflecting shRNA expression. All control tumours were GFP positive, whereas expression of GFP was barely detectable in tumours containing *shArk5*, and tumour relapse was accompanied by re-expression of ARK5 mRNA (Fig. 4f and Supplementary Fig. 11a). Acute depletion of ARK5 activated mTOR

and resulted in cell death and proliferative arrest *in vivo* (Supplementary Fig. 11b, c). Similar HCCs generated by co-expression of NRAS and MYC depended on ARK5, whereas HCCs generated by expression of NRAS in *Arf*^{-/-} cells did not (Supplementary Fig. 12).

Synthetic lethal interactions allow the elimination of tumour cells carrying genetic lesions that cannot be targeted using small molecules^{23–28}. We show that elevated energy consumption and addiction to mitochondrial glutaminolysis in cells expressing deregulated MYC establish a dependence on the kinase ARK5, which limits mTORC1 activity and maintains a high respiratory capacity. Our data support the view that oncogene-altered energy metabolism presents a new class of target molecules for tumour therapy.

METHODS SUMMARY

High-content screening. U2OS cells stably expressing MYC–ER under the control of a retroviral long terminal repeat (LTR) were transfected with siRNA in a 96-well format. Twenty-four hours after transfection, medium was replaced by medium with or without 200 nM OHT. Twenty-four hours later, cells were subjected to indirect immunofluorescence using cleaved-Parp (51-900017, BD) followed by Alexa488-conjugated secondary antibody (Invitrogen). To reduce potential off-target effects, a pool of four individual siRNAs was transfected for each kinase. For automated data acquisition, the BD Pathway 855 bioimager (BD Biosciences) was used.

Metabolomic analysis. Cells were incubated with ¹³C₆-glucose (Sigma) for 3 min or with ¹³C₅-glutamine (Sigma) for 7 min. After labelling, cells were quenched with MeOH/water (1:1; –20 °C), extracted with CHCl₃/MeOH/water and analysed by gas chromatography–time of flight–mass spectrometry (GC–TOF–MS) as described previously²⁹. Mass isotopomers were extracted from individual mass spectra using the MetMax software tool³⁰ and relative label incorporation was calculated after normalization to total area of measured analytes. For calculation of absolute concentrations of TCA cycle intermediates (all from Sigma) 8-point calibration curves spanning 2–3 orders of magnitudes were measured together with the samples.

Animal experiments. To deplete ARK5 in murine HCC cells, shRNAs were cloned into a retroviral vector for doxycycline-regulatable shRNA expression²². Early passage murine hepatoma cells were outgrown from genetically defined hepatocellular carcinomas (*Myc*; *Akt*; *p53*^{-/-}). Cells were transduced with *Ark5*-shRNA-expressing retroviruses at low multiplicity of infection to ensure single-copy retroviral integration. After puromycin selection, 1,000,000 cells were transplanted under the liver capsula of female C57BL/6 mice. shRNA expression was induced by administration of doxycycline via the drinking water (1 mg ml⁻¹ with 2% sucrose). Tumour progression was monitored by abdominal palpation and whole-body GFP imaging. All mice were maintained under pathogen-free conditions in accordance with the institutional guidelines of the Helmholtz Centre for Infection Research. All animal experiments were approved by the German legal authorities.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 June 2011; accepted 6 February 2012.

- Dang, C. V., Le, A. & Gao, P. MYC-induced cancer cell energy metabolism and therapeutic opportunities. *Clin. Cancer Res.* **15**, 6479–6483 (2009).
- Eilers, M. & Eisenman, R. N. Myc's broad reach. *Genes Dev.* **22**, 2755–2766 (2008).
- Lizcano, J. M. *et al.* LKB1 is a master kinase that activates 13 kinases of the AMPK subfamily, including MARK/PAR-1. *EMBO J.* **23**, 833–843 (2004).
- Suzuki, A. *et al.* ARK5 is a tumor invasion-associated factor downstream of Akt signaling. *Mol. Cell. Biol.* **24**, 3526–3535 (2004).
- Clark, K., Plater, L., Pegg, M. & Cohen, P. Use of the pharmacological inhibitor BX795 to study the regulation and physiological roles of TBK1 and I κ B kinase ϵ : a distinct upstream kinase mediates Ser-172 phosphorylation and activation. *J. Biol. Chem.* **284**, 14136–14146 (2009).
- Bissonnette, R. P., Echeverri, F., Mahboubi, A. & Green, D. R. Apoptotic cell death induced by *c-myc* is inhibited by *bcl-2*. *Nature* **359**, 552–554 (1992).
- Humbert, N. *et al.* Regulation of ploidy and senescence by the AMPK-related kinase NUAK1. *EMBO J.* **29**, 376–386 (2010).
- Shackelford, D. B. & Shaw, R. J. The LKB1–AMPK pathway: metabolism and growth control in tumour suppression. *Nature Rev. Cancer* **9**, 563–575 (2009).
- Hawley, S. A. *et al.* Use of cells expressing γ subunit variants to identify diverse mechanisms of AMPK activation. *Cell Metab.* **11**, 554–565 (2010).
- Oakhill, J. S. *et al.* β -Subunit myristoylation is the gatekeeper for initiating metabolic stress sensing by AMP-activated protein kinase (AMPK). *Proc. Natl Acad. Sci. USA* **107**, 19237–19241 (2010).

11. Choo, A. Y. *et al.* Glucose addiction of TSC null cells is caused by failed mTORC1-dependent balancing of metabolic demand with supply. *Mol. Cell* **38**, 487–499 (2010).
12. Carayol, N. *et al.* Critical roles for mTORC2- and rapamycin-insensitive mTORC1-complexes in growth and survival of BCR-ABL-expressing leukemic cells. *Proc. Natl Acad. Sci. USA* **107**, 12469–12474 (2010).
13. Um, S. H. *et al.* Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature* **431**, 200–205 (2004).
14. Gao, P. *et al.* c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458**, 762–765 (2009).
15. Yuneva, M., Zamboni, N., Oefner, P., Sachidanandam, R. & Lazebnik, Y. Deficiency in glutamine but not glucose induces MYC-dependent apoptosis in human cells. *J. Cell Biol.* **178**, 93–105 (2007).
16. Wise, D. R. *et al.* Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl Acad. Sci. USA* **105**, 18782–18787 (2008).
17. Kim, J., Kundu, M., Viollet, B. & Guan, K. L. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nature Cell Biol.* **13**, 132–141 (2011).
18. Zhang, H. *et al.* HIF-1 inhibits mitochondrial biogenesis and cellular respiration in VHL-deficient renal cell carcinoma by repression of C-MYC activity. *Cancer Cell* **11**, 407–420 (2007).
19. Li, F. *et al.* Myc stimulates nuclearly encoded mitochondrial genes and mitochondrial biogenesis. *Mol. Cell. Biol.* **25**, 6225–6234 (2005).
20. van de Wetering, M. *et al.* The β -catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
21. Zender, L. *et al.* An oncogenomics-based *in vivo* RNAi screen identifies tumor suppressors in liver cancer. *Cell* **135**, 852–864 (2008).
22. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnol.* **29**, 79–83 (2011).
23. Campaner, S. *et al.* Cdk2 suppresses cellular senescence induced by the *c-myc* oncogene. *Nature Cell Biol.* **12**, 54–59 (2010).
24. Scholl, C. *et al.* Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. *Cell* **137**, 821–834 (2009).
25. Barbie, D. A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108–112 (2009).
26. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
27. Luo, J. *et al.* A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* **137**, 835–848 (2009).
28. Yang, D. *et al.* Therapeutic potential of a synthetic lethal interaction between the MYC proto-oncogene and inhibition of aurora-B kinase. *Proc. Natl Acad. Sci. USA* **107**, 13836–13841 (2010).
29. Kempa, S. *et al.* A plastid-localized glycogen synthase kinase 3 modulates stress tolerance and carbohydrate metabolism. *Plant J.* **49**, 1076–1090 (2007).
30. Kempa, S. *et al.* An automated GC \times GC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential ^{13}C -labelling experiments: a case study for photoautotrophic-mixotrophic grown *Chlamydomonas reinhardtii* cells. *J. Basic Microbiol.* **49**, 82–91 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Baumann for immunohistochemical staining, B. Samans, A. Filmer and M. Krause for help with the siRNA screen and expression array analysis, A. Rath for help with analysis of human tumour samples, W. Kolch for siRNA targeting LATS1, S. Lowe and J. Zuber for providing the tGMP vector, S. Martin for His-tagged annexin V expression vector, P. Schirmacher for human HCC cell lines, V. Ellenrieder for human pancreatic cancer cell lines, H. Riedesel and the staff of the Helmholtz Centre for Infection Research animal facility for technical assistance. We thank P. Gallant and all members of the Eilers laboratory for critical reading of the manuscript. Funding for this work was provided by the GROWTHSTOP consortium of the European Union, the Deutsche Forschungsgemeinschaft via the Transregio 17 (M.E.), the Transregio 77 (L.Z.) and the Emmy Noether Programme ZE 545/2-1 (L.Z.), the Bundesministerium für Bildung und Forschung, the Senate of Berlin, the University of Würzburg Graduate School of Life Sciences and the “Rebirth” Cluster of Excellence.

Author Contributions L.L., J.U., J.M., T.R.K., N.M. and R.R. performed the experiments, R.M. analysed human tumour samples, and T.W. analysed the mouse model of HCC. L.A. and S.K. performed all metabolomic and proteomic analyses and questions regarding the technology should be addressed to S.K. L.R. performed bio-informatic analysis. L.Z., S.K., M.E. and D.J.M. designed experiments, and M.E. and D.J.M. wrote the paper.

Author Information Array results have been deposited in ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>) under accession number E-MEXP-3543. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.E. (martin.eilers@biozentrum.uni-wuerzburg.de) or D.J.M. (daniel.murphy@biozentrum.uni-wuerzburg.de).

METHODS

Cell culture. U2OS, MRC5, HCC4977 and Phoenix cells were cultured in DMEM (4.5 g ml⁻¹ glucose; 2 mM glutamine) containing 10% FBS, 100 U ml⁻¹ penicillin and 100 mg ml⁻¹ streptomycin (PAA). Where indicated, 200 nM 4-hydroxytamoxifen (OHT; Sigma), 5 µg ml⁻¹ doxycycline (Sigma), 2 mM caffeine (Sigma), 5 µg ml⁻¹ oligomycin (Sigma), 500 µM AICAR (Sigma), 100 nM rapamycin (LC Labs), 50 nM BX795 (Invitrogen) or 10 µM OSI-027 (Active Biochemicals) were added to the culture media. For ATP depletion experiments, cells were cultured in glucose- and glutamine-free DMEM (Sigma #D5030) supplemented with penicillin/streptomycin and 10% dialysed FBS (PAA). Apoptosis was measured by propidium iodide/annexin V (Molecular Probes) labelling and analysed using a BD FACS Canto II. Alternatively, cells were harvested by mild trypsinization and fixed in 80% ethanol, stained with 20 µg ml⁻¹ propidium iodide in 0.1% Triton X-100/PBS containing 0.2 mg ml⁻¹ RNAase A, and cellular DNA content was measured by FACS. siGENOME RNA oligonucleotides were purchased from Thermo Scientific Dharmacon, and cells were transfected using Dharmafect1 (Dharmacon) or Lipofectamine RNAiMAX (Invitrogen). Medium was replenished 24 h after transfection. Cellular ATP levels were determined using a luminescence-based ATP detection kit, ATPlite (PerkinElmer). For measurement of glutamine uptake, U2OS MYC-ER cells were incubated in DMEM medium and pulsed for 1 min with 1 µM ¹⁴C-labelled glutamine. The cells were washed three times with PBS and lysed in RIPA buffer for 30 min on ice. The lysates were analysed with a scintillation counter (Wallac 1410, Pharmacia). Mitochondrial membrane potential was measured FACS analysis of DiOC6(3) (Molecular Probes) stained cells, according to the manufacturer's instructions. Reactive oxygen species were likewise measured using CellROX Deep Red (Invitrogen). Oxygen consumption rates were calculated from parafilm-sealed 10-cm dishes of cells using a Clarke O₂ electrode.

High-content screening. U2OS cells stably expressing MYC-ER under the control of a retroviral long terminal repeat (LTR) were transfected with siRNA in a 96-well format. Twenty-four hours after transfection, medium was replaced by medium with or without 200 nM OHT. Twenty-four hours later, cells were subjected to indirect immunofluorescence using cleaved-Parp (51-9000017, BD) followed by Alexa488-conjugated secondary antibody (Invitrogen). To reduce potential off-target effects, a pool of four individual siRNAs was transfected for each kinase. For automated data acquisition, the BD Pathway 855 bioimager (BD Biosciences) was used.

Immunoblotting and antibodies. The following antibodies were purchased from Cell Signaling Technology: ARK5 (#4458), mTOR (#2972), phospho-mTOR (#2971), AMPK (#2532), phospho-AMPK^{Thr172} (#2535), S6 ribosomal protein (#2212), phospho-S6^{S240/244} (#2215), tuberin/TSC2 (D57A9, #3990), phospho-tuberin/TSC2^{Ser1387} (#5584), p70 S6 kinase (49D7, #2708), phospho-p70 S6K^{Thr389} (108D2, #9234). Anti-β-actin (AC-15, #A5441) and anti-mouse-ER (M-20) were purchased from Sigma. Cytochrome C Oxidase VIb (ab110266) and NduFS4 (ab87399) antibodies were purchased from Abcam. 9E10 and anti-mouse-ER were used to detect MYC and MYC-ER. Primary antibodies were used at 1:1,000, except for β-actin, which was used at 1:50,000. Secondary antibodies were purchased from Amersham.

Protein synthesis. For protein synthesis measurement U2OS cells were incubated for 2 h with ³H-leucine (50 µCi ml⁻¹ medium) in DMEM. The cells were washed twice with PBS, after which the cells were incubated in 10% TCA for 10 min on ice. The TCA treatment was repeated twice for 5 min each followed by a washing step in MeOH. The cells were air dried and lysed in 0.3 M NaOH, 1% SDS for 30 min at room temperature (20–22 °C). The lysates were mixed with scintillation fluid (Rotiszint exo plus, Carl Roth) and measured with a scintillation counter (Wallac 1410, Pharmacia).

Plasmids. To construct shRNAs for human ARK5 and AMPK, hairpin-encoding oligonucleotides were annealed and ligated into pRetroSuper vector. The following targeting sequences were used: ARK5-1, GATGACAACTGCAATATTA; ARK5-2 GGACAGTAATGATGTGATG; ARK5-3, AGGACAAAATTAAGGATGA; AMPK-1, AAGTCAAAGTCGACCAAAAT; AMPK-2, GCATAAAGTAGCTGTGAAG; AMPK-3, CAGCCGAGAAGCAGAAACA; AMPK-4, CCATACCCITGATGAATTA. ARK5-3 was used for most experiments. To deplete murine ARK5, hairpin encoding nucleotides targeting AGGGATTACTGGCATGGT (4977) and CGGTGGATGCTGATGGTGA (967) were inserted into pTGMP.

A full-length mouse cDNA encoding *Ark5* was obtained from ImaGenes, and subcloned into pBabe vector. Site-directed mutagenesis using the QuikChange XL kit (Stratagene) was performed to generate constructs expressing mutant mouse ARK5. The following primers were used: K85A, forward, 5'-CCGAGTGGTTGCTATAGCATCCATCCGTAAGGAC-3'; K85A, reverse, 5'-GTCCTTACGGATGGATGCTATAGCAACCACTCGG-3'; T212A, forward, 5'-CAGAAGGACAAGTCTTGCAAGCATTTTGTGGGAGCCCACTC-3'; T212A, reverse, 5'-GAGTGGGCTCCACAAAATGCTTGAAGAACTGTCTCTCTG-3'; S601A,

forward, 5'-GCCCGCCAGCGCATCCGCGCTTGGCTCTCTGCTGAAAC-3'; S601A, reverse, 5'-GTTTTCAGCAGAGACGCAAGCGCGGATGCGCTGGCGGGC-3'.

Human tissues. Paraffin blocks previously used for diagnostic purposes were taken from the files of the Institute of Pathology of the University of Marburg. Tissues were fixed with 10% formalin. In all, 26 cases of colorectal carcinoma, 20 cases of hepatocellular carcinoma (of which 11 were co-examined for MYC expression) and 21 cases of pancreatic ductal adenocarcinoma were examined. All tissues were pseudonymized, without use of any personal data of the patients, in accordance with the local Ethics Committee.

Animal experiments. To deplete ARK5 in murine HCC cells, shRNAs were cloned into a retroviral vector for doxycycline-regulatable shRNA expression²². Early passage murine hepatoma cells were outgrown from genetically defined hepatocellular carcinomas (*Myc; Akt; p53*^{-/-}). Cells were transduced with *Ark5*-shRNA-expressing retroviruses at low multiplicity of infection to ensure single-copy retroviral integration. After puromycin selection, 1,000,000 cells were transplanted under the liver capsule of female C57BL/6 mice. shRNA expression was induced by administration of doxycycline via the drinking water (1 mg ml⁻¹ with 2% sucrose). Tumour progression was monitored by abdominal palpation and whole-body GFP imaging. All mice were maintained under pathogen-free conditions in accordance with the institutional guidelines of the Helmholtz Centre for Infection Research. All animal experiments were approved by the German legal authorities.

Immunohistochemistry. 3–4-µm-thick paraffin sections were mounted on poly-L-lysine-coated slides, incubated at 58 °C, and deparaffinized. For antigen retrieval, sections were incubated in an antigen retrieval buffer (Tris-EDTA buffer, pH 9.0; Dako) for 30 min in a household steamer. The following incubations, including prior blocking of endogenous peroxidase activity, were performed using an automated immunohistochemistry apparatus (Autostainer plus; Dako): anti-ARK5 (Cell Signaling Technology, #4458) and anti c-MYC (Epitomics, #1472-1) antibodies were used at a dilution of 1:100. Detection was via Dako REAL Detection System Peroxidase/DAB+, rabbit/mouse; Dako) followed by staining with 3,3'-diaminobenzidine (DAB). For mild counterstaining, Mayer's haematoxylin solution was used. Where indicated, sections of snap frozen tumour tissues were subjected to TdT-mediated dUTP nick end labelling (TUNEL) staining (Roche) and to Ki-67 (Dianova) immunohistochemistry using standard protocols. For negative controls, the primary antibody was replaced by buffer or an irrelevant monoclonal antibody. For phospho-S6, 5-µm sections were subjected to heat-based antigen retrieval in 10 mM sodium citrate, blocked for 1 h in 3% BSA and incubated with primary antibody (Cell Signaling #5364) at 1:5,000 overnight at 4 °C. Detection was with signal stain boost (Cell Signaling) for 45 min at room temperature.

Metabolomic analysis. Cells were incubated with ¹³C₆-glucose (Sigma) for 3 min or with ¹³C₅-glutamine (Sigma) for 7 min. After labelling, cells were quenched with MeOH/water (1:1; -20 °C), extracted with CHCl₃/MeOH/water and analysed by gas chromatography–time of flight–mass spectrometry (GC-TOF-MS) as described previously²⁹. Mass isotopomers were extracted from individual mass spectra using the MetMax software tool³⁰ and relative label incorporation was calculated after normalization to total area of measured analytes. For calculation of absolute concentrations of TCA cycle intermediates (all from Sigma) 8-point calibration curves spanning 2–3 orders of magnitudes were measured together with the samples.

Proteome analysis. To achieve quantitative proteome data a 'heavy' proteome reference was spiked in equal amounts in every sample. HCC cells were grown in SILAC medium for 5 passages. SILAC medium was prepared as described previously³¹. In essence, DMEM lacking arginine and lysine was supplemented with 10% dialysed FBS (Sigma, F0392) and antibiotics. Amino acids (84 mg l⁻¹ ¹³C₆ ¹⁵N₄ L-arginine plus 146 mg l⁻¹ ¹³C₆ ¹⁵N₂ L-lysine) were added to obtain 'heavy' medium. Harvested cells were lysed in an appropriate amount of urea buffer (8 M urea, 50 mM Tris-HCl, pH 7.4). The lysates were cleared by centrifugation at 14,000 r.p.m. for 15 min at 4 °C. Disulphide bridges were then reduced in DTT 2 mM for 30 min at 25 °C and successively free cysteines were alkylated in 11 mM iodoacetamide for 20 min at room temperature in the darkness. LysC digestion was performed by adding LysC (Wako) in a ratio 1:40 (w/w) to the sample and incubating it for 18 h under gentle shaking at 30 °C. After LysC digestion, the samples were diluted 3 times with 50 mM ammonium bicarbonate solution, 7 µl of immobilized trypsin (Applied Biosystems) was added and samples were incubated for 4 h under rotation at 30 °C. Digestion was stopped by acidification with 10 µl of trifluoroacetic acid and removal of trypsin beads by centrifugation. After digestion peptides were extracted and desalted before analysis by mass spectrometry. 5 µl were injected in duplicate on a LC-MS/MS system (Agilent 1200 (Agilent Technologies) and LTQ-Orbitrap Velos (Thermo)), using a 240 min gradient ranging from 5% to 40% of solvent B (80% acetonitrile, 0.1% formic acid; solvent A = 5% acetonitrile, 0.1% formic acid). For the chromatographic separation, a ~20-cm-long capillary (75 µm inner diameter) was packed with 3 µm C18 beads (ReprosilPur C18 AQ, Dr. Maisch). On one end of the

capillary a nanospray tip was generated using a laser puller (P-2000 Laser Based Micropipette Puller, Sutter Instruments), allowing fretless packing. The nanospray source was operated with a spray voltage of 1.9 kV and an ion transfer tube temperature of 260 °C. Data were acquired in data dependent mode, with one survey MS scan in the Orbitrap mass analyser (resolution 60,000 at m/z 400) followed by up to 20 MS/MS scans in the ion trap on the most intense ions (intensity threshold, 500 counts). Once selected for fragmentation, ions were excluded from further selection for 30 s, to increase new sequencing events.

Data analysis. Raw data were analysed using the MaxQuant proteomics pipeline (v.1.1.1.36) and the built in the Andromeda search engine³² with the International Protein Index Mouse database. Carbamidomethylation of cysteines was chosen as fixed modification, oxidation of methionine and acetylation of N terminus were chosen as variable modifications. The search engine peptide assignments were filtered at 1% FDR and the feature match between runs was not enabled; second

peptide feature was enabled, while other parameters were left as default. For SILAC samples, two ratio counts were set as threshold for quantification. Data analysis was performed using custom tools in Microsoft Excel and R. Gene Ontology analysis was performed using David tool³³.

Microarray experiments were performed using Agilent whole human genome microarray kit (G4112F; Agilent Technologies) according to the manufacturer's instructions.

31. Ong, S. E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols* **1**, 2650–2660 (2007).
32. Cox, J. *et al.* Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
33. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response

Zhao Chen^{1,2,3}, Katherine Cheng^{2,3}, Zandra Walton^{2,3}, Yuchuan Wang^{4,5,6}, Hiromichi Ebi^{1,7}, Takeshi Shimamura⁸, Yan Liu^{1,2,3}, Tanya Tupper⁴, Jing Ouyang², Jie Li⁹, Peng Gao^{2,3}, Michele S. Woo², Chunxiao Xu^{1,2,3}, Masahiko Yanagita², Abigail Altabef², Shumei Wang¹⁰, Charles Lee¹⁰, Yuji Nakada¹¹, Christopher G. Peña¹¹, Yanping Sun^{4,5}, Yoko Franchetti¹², Catherine Yao², Amy Saur⁴, Michael D. Cameron¹³, Mizuki Nishino^{5,6}, D. Neil Hayes¹⁴, Matthew D. Wilkerson¹⁴, Patrick J. Roberts¹⁴, Carrie B. Lee¹⁴, Nabeel Bardeesy⁷, Mohit Butaney², Lucian R. Chirieac¹⁰, Daniel B. Costa¹⁵, David Jackman², Norman E. Sharpless¹⁴, Diego H. Castrillon¹¹, George D. Demetri³, Pasi A. Jänne^{1,2,16}, Pier Paolo Pandolfi¹⁷, Lewis C. Cantley^{18,19}, Andrew L. Kung^{4,20}, Jeffrey A. Engelman^{1,7} & Kwok-Kin Wong^{1,2,3,16}

Targeted therapies have demonstrated efficacy against specific subsets of molecularly defined cancers^{1–4}. Although most patients with lung cancer are stratified according to a single oncogenic driver, cancers harbouring identical activating genetic mutations show large variations in their responses to the same targeted therapy^{1,3}. The biology underlying this heterogeneity is not well understood, and the impact of co-existing genetic mutations, especially the loss of tumour suppressors^{5–9}, has not been fully explored. Here we use genetically engineered mouse models to conduct a ‘co-clinical’ trial that mirrors an ongoing human clinical trial in patients with *KRAS*-mutant lung cancers. This trial aims to determine if the MEK inhibitor selumetinib (AZD6244)¹⁰ increases the efficacy of docetaxel, a standard of care chemotherapy. Our studies demonstrate that concomitant loss of either *p53* (also known as *Tp53*) or *Lkb1* (also known as *Stk11*), two clinically relevant tumour suppressors^{6,9,11,12}, markedly impaired the response of *Kras*-mutant cancers to docetaxel monotherapy. We observed that the addition of selumetinib provided substantial benefit for mice with lung cancer caused by *Kras* and *Kras* and *p53* mutations, but mice with *Kras* and *Lkb1* mutations had primary resistance to this combination therapy. Pharmacodynamic studies, including positron-emission tomography (PET) and computed tomography (CT), identified biological markers in mice and patients that provide a rationale for the differential efficacy of these therapies in the different genotypes. These co-clinical results identify predictive genetic biomarkers that should be validated by interrogating samples from patients enrolled on the concurrent clinical trial. These studies also highlight the rationale for synchronous co-clinical trials, not only to anticipate the results of ongoing human clinical trials, but also to generate clinically relevant hypotheses that can inform the analysis and design of human studies.

Activating *KRAS* mutations are found in 15–30% of all patients with non-small cell lung cancer (NSCLC), and predict poor outcome in response to conventional treatment regimens^{13,14}. Preclinical studies have suggested that inhibition of MAPK/ERK kinase (MEK) may be effective against *KRAS*-mutant NSCLC¹⁵, prompting an ongoing

human clinical trial comparing docetaxel monotherapy (standard of care) to docetaxel combined with the MEK inhibitor selumetinib (AZD6244). Although the sole genetic entry criteria for patients on this trial is the presence of *KRAS* mutations, the complexity of NSCLC dictates that many tumours will harbour concomitant genetic alterations that may modulate response to therapy. To mirror this human clinical trial in a murine co-clinical trial, and to investigate the modulating effects of concomitant tumour suppressor loss, we generated cohorts of genetically engineered mice with either *Kras*, *Kras* and *p53* (*Kras/p53*) or *Kras/Lkb1* mutant lung cancers. Activation of *Kras*(G12D) and inactivation of *p53* or *Lkb1* in the lung epithelium was achieved using nasal instillation of adenovirus encoding the CRE recombinase⁷. Mice with established disease, defined by tachypnoea, hypoxaemia on pulse oximetry¹⁶, and bulk disease on magnetic resonance imaging (MRI)¹⁷, were randomized to receive either docetaxel 16 mg kg^{−1} every other day by intraperitoneal injection¹⁸ (Supplementary Table 1), selumetinib at 25 mg kg^{−1} daily by oral gavage¹⁹, or docetaxel in combination with selumetinib. Treatment response was determined by serial MRI. Tumour volumes were reconstructed from the MRI images (Supplementary Fig. 1a) with a high level of inter-operator reliability (Supplementary Fig. 1b; 95% confidence interval, −25.6% to +31.4%). On the basis of these performance metrics, and paralleling human response criteria, we used a threshold of 30% change in tumour volume to define progressive disease and partial response.

For tumours with only *Kras* mutation, treatment with docetaxel monotherapy resulted in a modest rate of response, with 30% of mice achieving a partial response (Fig. 1a, c). Mice bearing *Kras* tumours with concurrent loss of *p53* or *Lkb1* had markedly lower response rates to docetaxel treatment (5% and 0%, respectively), and more of these animals demonstrated progressive disease on MRI or progression to morbidity (Fig. 1a and Supplementary Table 2). The addition of selumetinib to docetaxel treatment provided substantial benefit (Fig. 1b, c), with the overall response rate increased to 92% in *Kras*-mutant cancers ($P = 2.8 \times 10^{-5}$, Fisher exact test compared to docetaxel alone) and 61% in *Kras/p53* mice ($P = 2.7 \times 10^{-4}$). In contrast, for

¹Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. ²Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ³Ludwig Center at Dana-Farber/Harvard Cancer Center, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁴Lurie Family Imaging Center, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁵Department of Imaging, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁶Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁷Department of Medical Oncology, Massachusetts General Hospital Cancer Center, Boston, Massachusetts 02114, USA. ⁸Department of Molecular Pharmacology and Therapeutics, Oncology Institute, Loyola University Chicago Stritch School of Medicine, Maywood, Illinois 60153, USA. ⁹Center for Survey Statistics & Methodology, Department of Statistics & Statistical Laboratory, Iowa State University, Ames, Iowa 50010, USA. ¹⁰Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ¹¹Department of Pathology and Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, Texas 75390, USA. ¹²Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA. ¹³Translational Research Institute, The Scripps Research Institute, Jupiter, Florida 33458, USA. ¹⁴The Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599, USA. ¹⁵Division of Hematology/Oncology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁶Low Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. ¹⁷Cancer Genetics Program, Beth Israel Deaconess Cancer Center, Department of Medicine and Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁸Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁹Division of Signal Transduction, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02115, USA. ²⁰Department of Pediatric Oncology, Dana Farber Cancer Institute and Children's Hospital, Boston, Massachusetts 02115, USA.

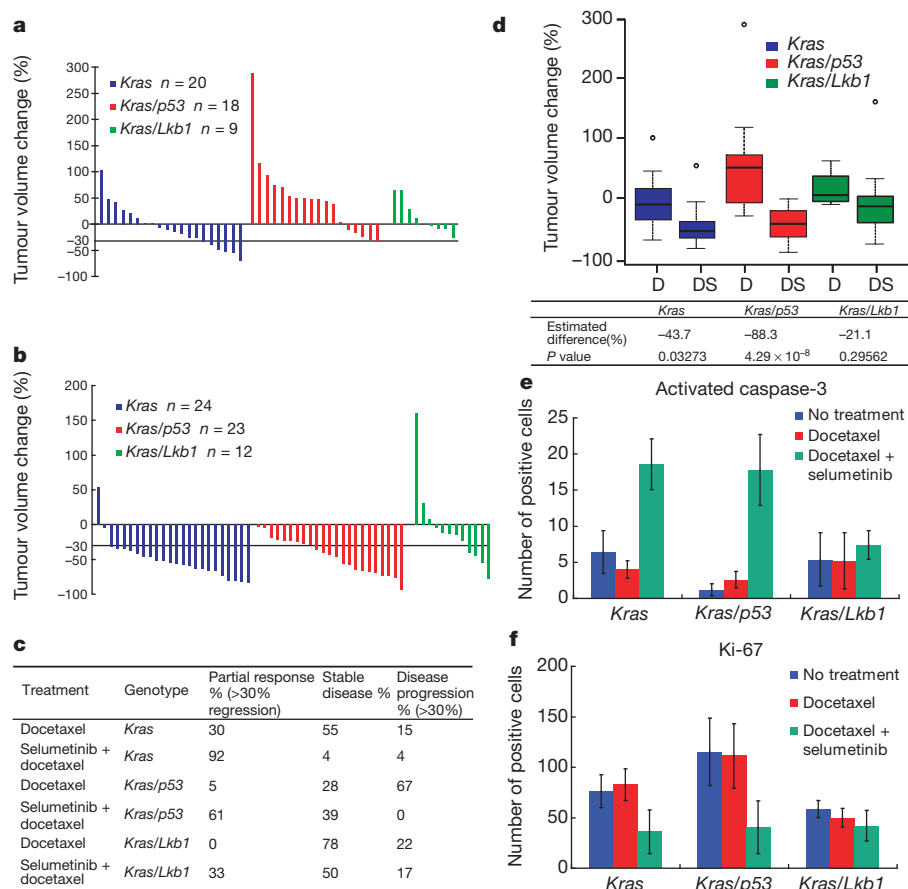


Figure 1 | Docetaxel and selumetinib combination therapy is more efficacious than docetaxel monotherapy in *Kras* and *Kras/p53* lung cancers. **a**, Waterfall plot showing tumour response after 2 weeks of docetaxel treatment at 16 mg kg^{-1} every 2 days. Each column represents one individual mouse, with data expressed relative to the pre-treatment tumour volume. **b**, Waterfall plot showing tumour response after 2 weeks of docetaxel treatment at 16 mg kg^{-1} every 2 days in combination with daily selumetinib at 25 mg kg^{-1} . **c**, Response rate of docetaxel and selumetinib combination therapy and docetaxel only in mice bearing tumours with different genotypes. **d**, Box plot showing tumour response for different genotypes with either docetaxel monotherapy (D) or combination treatment (DS). Lines depict median response, small circles indicate outliers. Estimated magnitude of difference between single and combination treatment within each genotype and corresponding one-sided *P* values obtained by likelihood ratio test. **e**, **f**, Number of activated caspase-3 (**e**) and Ki-67 (**f**) positive cells per microscopic field in mice of different genotypes after short-term treatment. Data represent the average of 5 different fields \pm standard deviation (s.d.) from 1–3 different mice (see Supplementary Table 6 for detailed information).

Kras/Lkb1 mutant cancers the addition of selumetinib to docetaxel led to only a modest improvement in overall response, with 33% of the mice achieving a partial response (Fig. 1b, c). Compared to the other genotypes, *Kras/Lkb1* mice had a significantly lower rate of response to combined treatment with selumetinib and docetaxel ($P = 0.0009$, 3×2 contingency Fisher exact test).

The magnitude of change in volume confirmed that tumours with *Kras* or *Kras/p53* mutations were considerably more responsive to combination therapy compared to docetaxel alone. In contrast, the addition of selumetinib did not significantly reduce the volume of tumours with compound *Kras/Lkb1* mutations (Fig. 1d and Supplementary Fig. 1c, d). Concordantly, histopathological assessments of tumours collected after two doses of treatment revealed that the combination treatment increased apoptosis (Fig. 1e and Supplementary Fig. 2a) and reduced proliferation (Fig. 1f and Supplementary Fig. 2b) in the *Kras* and *Kras/p53* tumours compared to docetaxel alone, but this was not observed in *Kras/Lkb1* tumours (Fig. 1e, f and Supplementary Fig. 2a, b). These results demonstrate that combined treatment with selumetinib and docetaxel induces apoptosis and decreases proliferation in *Kras* and *Kras/p53* tumours, leading to antitumour efficacy, but that concomitant mutation of *Lkb1* confers primary resistance to the combination treatment.

Because repeated tumour biopsies are difficult in patients, we explored the use of ^{18}F -fluoro-2-deoxy-glucose PET (FDG-PET) as an early response indicator that could be used in the clinic. Comparison of FDG avidity, quantified by standardized uptake value (SUV)^{20,21} in lung cancers across the three different genotypes showed an overall higher FDG uptake in both *Kras/p53* and *Kras/Lkb1* tumours compared to *Kras* tumours (Fig. 2a; $P = 0.02$, one-way ANOVA). Expression of the glucose transporter GLUT1 (also known as SLC2A1) was elevated in *Kras/Lkb1* mutant tumours (Supplementary Fig. 3a), consistent with the increased baseline FDG-PET signal. To determine if this finding was applicable to human patients, we determined the pre-treatment

FDG avidity in nine patients with *KRAS*-mutated lung cancer. Tumours from three patients positive for LKB1 immunostaining had a mean maximum SUV (SUV_{max}) of 2.33, whereas tumours from six patients negative for LKB1 immunostaining had a mean SUV_{max} of 8.75 (Fig. 2b; $P = 0.048$, two-sided Wilcoxon).

We next used FDG-PET to assess early tumour metabolic changes after initiation of therapy. Treatment with docetaxel alone did not result in significant changes in tumour hypermetabolism in *Kras*-, *Kras/p53*- or *Kras/Lkb1*-tumour-bearing mice (Fig. 2c, d). Of note, some of the murine *Kras* lung cancer nodules were not FDG avid (Fig. 2a) and these were the most sensitive to single-agent docetaxel (data not shown). In contrast, within 24 h of the first dose of treatment with docetaxel and selumetinib, tumour hypermetabolism was markedly suppressed in both *Kras* and *Kras/p53* mice (Fig. 2c, d). However, *Kras/Lkb1*-mutant tumours had no appreciable decrease in FDG avidity when treated with the combination (Fig. 2c, d). Together, these results demonstrate that early changes in tumour metabolism measured by FDG-PET (Fig. 2c, d) are concordant with histopathological analysis of apoptosis and proliferation (Fig. 1e, f) and predict antitumour efficacy (Fig. 1a–c) of docetaxel and selumetinib in treating *Kras*-mutant lung cancers.

To assess the pharmacodynamic effects of treatment on the MEK–ERK signalling axis, we assayed pathway activation using phospho-ERK immunostaining of lung cancer nodules. At baseline, the ERK pathway was most activated in *Kras/p53*-mutant tumours (Fig. 3a, b). We observed substantially less phospho-ERK staining in *Kras/Lkb1* tumours, suggesting that the MEK–ERK pathway is not highly activated in these cancers. Treatment with docetaxel did not alter phospho-ERK staining, but, as expected, the addition of selumetinib decreased MEK–ERK signalling in the *Kras* and *Kras/p53* tumours (Fig. 3a, b).

We further evaluated cellular signalling from short-term-treated lung cancer nodules by immunoblotting tumour lysates. Concordant with immunostaining (Fig. 3a), elevated phospho-ERK and phospho-90RSK were observed in *Kras/p53* tumours relative to the other genotypes

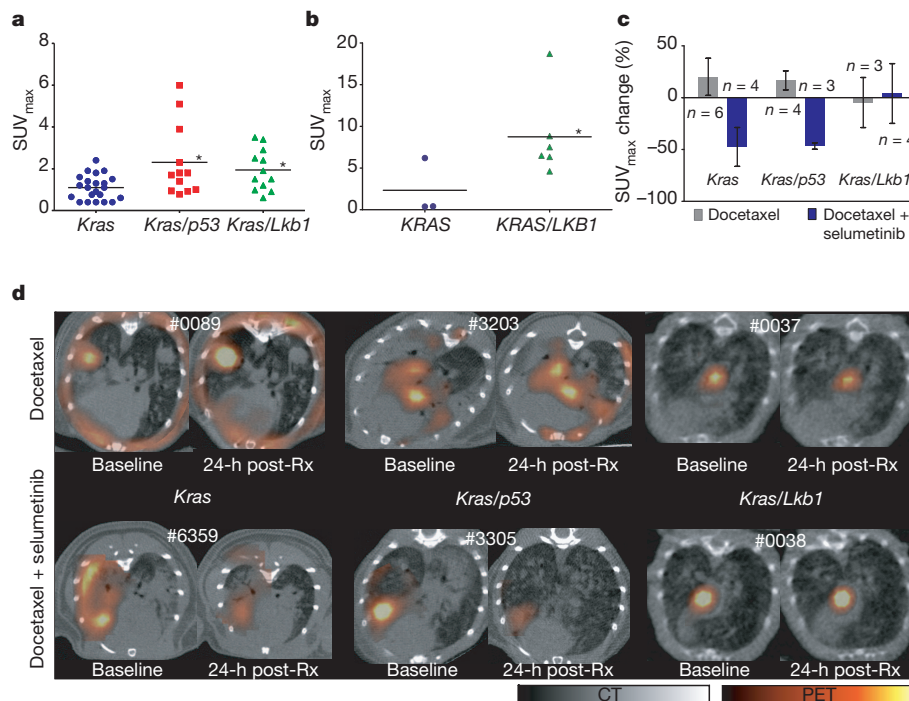


Figure 2 | FDG-PET predicts treatment response. **a**, FDG-PET signal intensity (SUV_{max}) in *Kras*, *Kras/p53* and *Kras/Lkb1* mutant mice. Statistical significance determined by rank sum test, with $*P < 0.05$ for *Kras* compared to *Kras/p53* mutant mice ($P = 0.019$), and *Kras* compared to *Kras/Lkb1* mutant mice ($P = 0.014$). **b**, FDG-PET signal intensity in patients with *KRAS* or *KRAS/LKB1* mutant tumours. Statistical significance determined by two-sided Wilcoxon with $*P = 0.048$. **c**, Comparisons of changes in FDG uptake by PET imaging after 1 day of treatment. Data are represented as mean \pm s.d. **d**, Representative FDG-PET/CT images of mice from different genotypes at baseline and 1 day after initiation of treatment. For each animal, the baseline and post-treatment (post-Rx) PET images are depicted with identical scales. The pseudocoloured FDG-PET images are fused with the grey-scale cross-sectional CT images.

(Fig. 3c and Supplementary Fig. 4a). *Kras/Lkb1* tumours displayed low basal activation of the MEK–ERK pathway (Fig. 3c and Supplementary Fig. 4a), consistent with immunostaining (Fig. 3a). Treatment with docetaxel had no discernable impact on the MEK–ERK pathway in any genotype (Fig. 3a–c and Supplementary Fig. 4b). Although selumetinib alone resulted in decreased phospho-ERK, residual activity was still present (Fig. 3c and Supplementary Fig. 4b). Treatment with both docetaxel and selumetinib more effectively eradicated phospho-ERK activity (Fig. 3c and Supplementary Fig. 4b). Pharmacokinetic studies suggested that selumetinib levels were elevated in the serum

and tumours of mice treated with selumetinib combined with docetaxel compared to selumetinib alone (Supplementary Table 3), perhaps providing a mechanism for the more potent suppression of MEK–ERK signalling by the combination (Fig. 3c). The potential relevance of these findings to human disease was investigated by assessing phospho-ERK staining in a set of 57 human NSCLC tumour samples with known *KRAS*, *p53* and *LKB1* mutation status. Consistent with our findings in murine tumours, of seven patients harbouring the *KRAS* activating mutation, the three patients with concurrent *p53* loss showed higher phospho-ERK activity (Fig. 3d).

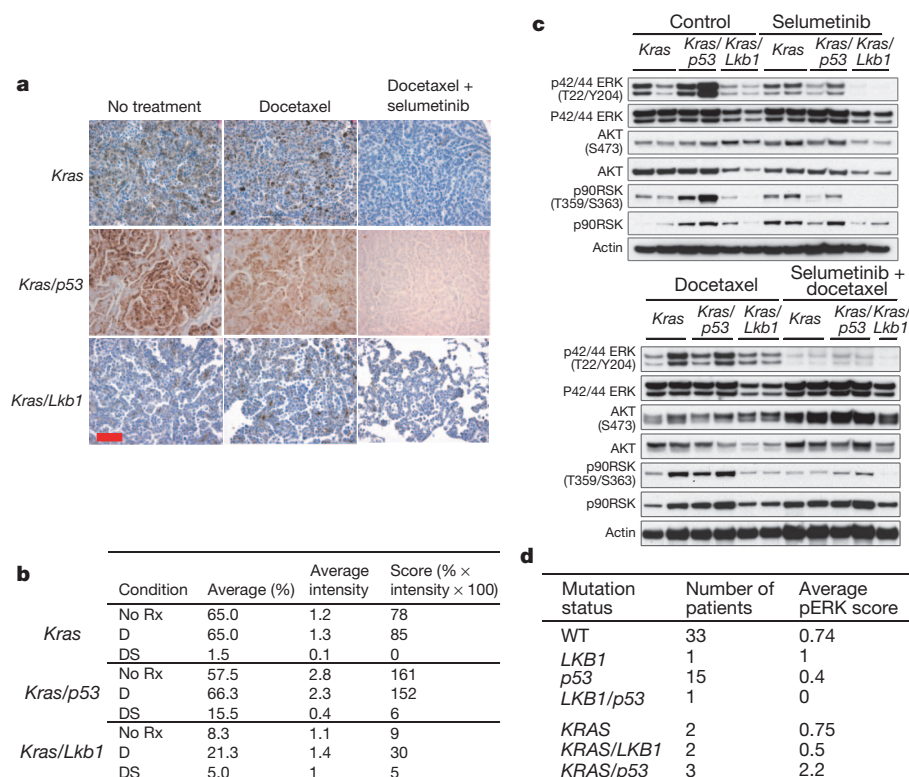


Figure 3 | Modulation of the MEK–ERK pathway in response to treatment is different across the three genotypes. **a**, Immunostaining of phospho-ERK before and after treatment with docetaxel alone or in combination with selumetinib. Scale bar, 50 μ m. **b**, Pathology score of phospho-ERK staining of mouse tumours shown in **a**. Both percentage of phospho-ERK positive cells and average intensity of phospho-ERK staining were scored for individual nodules, with a composite score derived by multiplying phospho-ERK positive percentage and average intensity. All samples were stained in the same batch. D, docetaxel; DS, docetaxel plus selumetinib; No Rx, untreated. **c**, Mice were subjected to treatment (two doses in 24 h) as indicated, and killed 3 h after the second dose. Western blot was used to analyse tumour lysates with the indicated antibodies. **d**, Human NSCLC patients grouped by mutation status as indicated in the first column. Mean phospho-ERK (pERK) score from immunostaining is shown for each subset.

The decreased activation of ERK phosphorylation in *Kras/Lkb1* tumours suggests that the proliferation of these tumours may be driven through other signalling pathways. On the basis of our prior studies^{19,22}, we investigated the activity of AKT and SRC in *Kras/Lkb1* mutant tumours. Immunoblotting with activation-state-specific antibodies revealed that *Kras/Lkb1*-mutant tumours have heightened activation of both AKT and SRC (Supplementary Fig. 3a, b), consistent with the finding of increased FDG avidity in *Kras/Lkb1* tumours (Fig. 2a, b), because PI3K regulates expression of GLUT1 (Supplementary Fig. 3a). These results suggest that concomitant mutation of *Kras* and *Lkb1* may alter the signalling circuitry in tumour cells from one dependent on MEK–ERK (in *Kras* and *Kras/p53* tumours) to one that has more active AKT and SRC pathways, resulting in primary resistance to docetaxel and selumetinib.

The concurrent human clinical trial does not include a treatment arm in which patients are treated with selumetinib alone, based on lack of efficacy in a phase II clinical trial in patients with NSCLC²³, and on our preclinical data in *Kras* genetically engineered mice¹⁹. In mice with *Kras*, *Kras/p53* and *Kras/Lkb1* tumours, treatment with selumetinib alone resulted in a heterogeneous reduction in FDG-PET uptake (Supplementary Fig. 5a), consistent with pharmacodynamic evidence that selumetinib alone partially attenuates MEK–ERK signalling within tumours (Fig. 3c and Supplementary Fig. 7c). However, no partial responses were achieved in any genotype with selumetinib monotherapy, although there was attenuation of tumour growth compared to untreated controls (Supplementary Fig. 5b). Together, these data suggest that selumetinib as monotherapy modulates MEK–ERK signalling in *Kras*-driven tumours, but is insufficient for clinical benefit in mice and humans.

We determined the long-term benefit of combined treatment with docetaxel and selumetinib in the *Kras*- and *Kras/p53*-mutant mice compared to chronic treatment with docetaxel monotherapy. We did not assess long-term treatment outcome in *Kras/Lkb1* animals given the primary resistance to both treatments in these animals (Figs 1–3). In mice with *Kras* tumours, treatment with docetaxel alone stabilized disease for several weeks, whereas the addition of selumetinib caused frank tumour regression and slower tumour re-growth (Fig. 4a and Supplementary Fig. 6a, b). Accordingly, the addition of selumetinib to docetaxel significantly prolonged progression-free survival (Fig. 4b). In mice with *Kras/p53* tumours, treatment with docetaxel alone largely resulted in progressive disease, whereas animals treated with a combination of docetaxel and selumetinib had initial disease regression before progression (Fig. 4a and Supplementary Fig. 6c), resulting in prolonged progression-free survival (Fig. 4c). These results demonstrate that the enhanced response to treatment with combined therapy translates to improved progression-free survival, albeit not outright cure, in mice bearing *Kras*- and *Kras/p53*-mutant tumours.

To investigate mechanisms of resistance upon disease progression, tumour nodules were isolated from moribund animals after long-term treatment with docetaxel and selumetinib. In all animals examined (5/5 in *Kras/p53* and 11/11 in *Kras*), tumour nodules showed recrudescence of ERK phosphorylation (Fig. 4d and Supplementary Fig. 7a), suggesting that acquired resistance could be partly due to reactivation of MEK–ERK signalling despite ongoing treatment with selumetinib. We evaluated treatment-resistant nodules for ERK amplification (Supplementary Fig. 7b), activation of parallel signalling pathways (Supplementary Fig. 7c), and drug pharmacokinetics (Supplementary Fig. 7d), and did not find consistent changes, suggesting more than one mechanism for pathway reactivation. Efforts to identify the diversity of mechanisms responsible for acquired resistance are ongoing.

This co-clinical study provides several insights and predictions that affect the interpretation of the concurrent human clinical trial. First, these results predict that combination therapy with docetaxel and selumetinib will be more effective than docetaxel alone in several subclasses of *KRAS*-mutant NSCLC. These data are consistent with the results of the human phase II clinical trial described in a recent

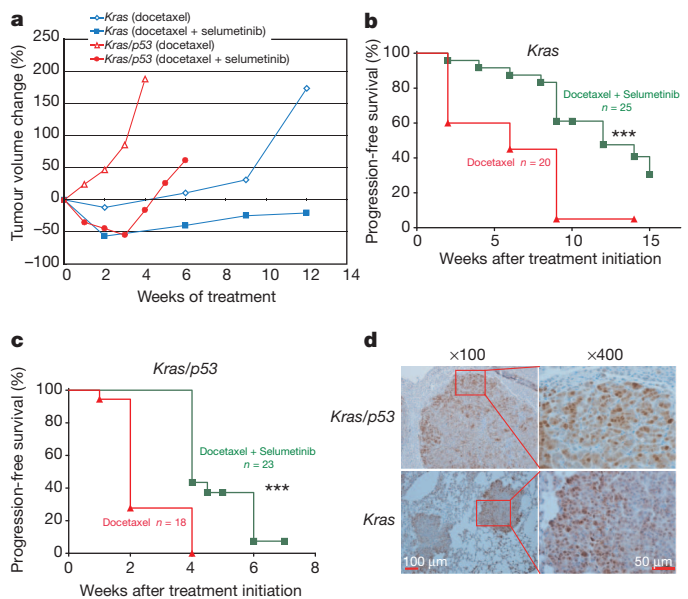


Figure 4 | Long-term treatment outcome in *Kras* and *Kras/p53* mice.

a, Tumour volume was longitudinally assessed by MRI imaging in *Kras* and *Kras/p53* mice treated with either docetaxel or docetaxel plus selumetinib. Data points represent median tumour volume relative to start of treatment for all available data at the indicated time point. **b**, Progression-free survival for *Kras* mice treated with either docetaxel or docetaxel plus selumetinib. Median survival for single and combination treatments was 6 weeks and 12 weeks respectively, with ***P = 0.0003 by log-rank test. **c**, Progression-free survival for *Kras/p53* mice treated with either docetaxel or docetaxel plus selumetinib. Median survival for single and combination treatments was 2 weeks and 4 weeks, respectively, with ***P < 0.0001 by log-rank test. Progression was defined as the time point when total tumour volume exceeded the baseline volume. **d**, Immunostaining of activation-specific phospho-ERK of tumours from *Kras/p53* and *Kras* mice with acquired resistance to docetaxel and selumetinib treatment.

press release (<http://phx.corporate-ir.net/phoenix.zhtml?c=123810&p=irol-newsArticle&ID=1611800>). However, our studies predict that concurrent mutation of *LKB1* will confer primary resistance to combination therapy, possibly through activation of parallel signalling pathways such as AKT and SRC. As *LKB1* status is not being prospectively assessed in the ongoing human clinical trial, inclusion of patients with cancers harbouring concurrent *LKB1* mutations may blunt differences between treatment arms based solely on *KRAS* status. These results suggest that a retrospective analysis of *p53* and *LKB1* status in samples from the concurrent human clinical trial is warranted, and lead us to advocate prospective analysis in future clinical trials with sufficient enrolment to all strata to enable sufficiently powered sub-group analyses.

Beyond assessing genetic modulators, co-clinical studies allow for validation of biomarker strategies and discovery of mechanisms of resistance that may benefit future clinical trials. In this study, we observed the potential utility of FDG-PET imaging as a biomarker strategy for identifying an enriched responder population and predicting long-term outcome. Prior studies have suggested that lung tumour hypermetabolism at the time of diagnosis predicts poor outcome in response to conventional therapies^{20,24}, our current study further suggests that high baseline FDG avidity may predict poor response to targeted therapy with selumetinib combined with docetaxel. Specifically, loss of *LKB1* function appears to confer increased FDG avidity, probably through upregulated expression of glucose transporters. As current approaches for assessing *LKB1* status are not comprehensive, FDG-PET imaging may represent a practicable patient stratification strategy. Furthermore, the current preclinical study suggests a role for repeat FDG-PET imaging early in the course of treatment as a potential predictor of outcome, as metabolic changes may be apparent within 24 h of initiating therapy. In these studies, we also observed reactivation of the MEK–ERK

signalling pathway in mice that became resistant to the combination of selumetinib and docetaxel. Although the exact mechanisms responsible for pathway reactivation remain to be elucidated, mechanisms of resistance discovered in co-clinical studies should be confirmed in human clinical trials by examining biopsy samples from patients who relapse on therapy. The ability to assess mechanisms of resistance in the preclinical setting may uncover rational combinatorial strategies that can be implemented in future clinical studies.

Building upon prior success using genetically engineered mouse models^{25–27}, the current study demonstrates that co-clinical trials can provide data that has value beyond predicting the outcome of clinical trials, and can rapidly generate new clinically relevant hypotheses that can affect how the concurrent human clinical trial is analysed, and inform the design of future clinical studies. As similar efforts are undertaken in other cancer disease types, we anticipate that murine co-clinical trials will enable more effective oncology drug development.

METHODS SUMMARY

Mice. Mouse strains harbouring a conditional activating mutation (G12D) at the endogenous *Kras* locus, conditional *Lkb1* knockout, and conditional *p53* knockout were described previously⁷. Genotypes were confirmed by PCR (Supplementary Fig. 8). All studies were performed on protocols approved by Dana-Farber Cancer Institute and University of North Carolina Animal Care and Use Committees, and all mice used are listed in Supplementary Table 4.

MRI quantification. 3D Slicer was used to reconstruct MRI volumetric measurements^{28,29} (Supplementary Fig. 1a). To assess variation between independent operators, Bland–Altman analysis was performed using quantification results from the two operators on a total of 16 MRI scan images (Supplementary Fig. 1b).

Pharmacokinetics. Docetaxel concentrations in serum, lung, and tumour nodules was determined using published doses (8 and 16 mg kg^{−1}). Drug concentration was determined 3 h after the last dose (Supplementary Table 1). All serum concentrations were within the range found in the clinical setting³⁰. Selumetinib was administered as previously described¹⁹, and pharmacokinetics in mice were also documented (Supplementary Table 3).

PET/CT studies. All murine FDG-PET/CT studies were performed with a pre-clinical small animal PET/CT system (Siemens Inveon) after injection with 14 MBq of ¹⁸F-FDG. Mice used for PET/CT studies are listed in Supplementary Table 5.

Human samples and clinical information. All human samples and clinical information were obtained under Institutional Review Board approved protocols (02-180 and 07-0120), and patient information is listed in Supplementary Tables 7 and 8. A tissue microarray (TMA) was generated from genotyped human lung cancer samples as previously described⁷. The TMA was immunostained for phospho-ERK and scored by a pathologist blinded to patient information.

Immunohistochemistry staining. Immunohistochemical analyses assessing phospho-ERK, activated caspase 3, and Ki-67 were performed as previously described¹⁹. Scoring was done by a pathologist, using the same parameters used for scoring human specimens.

Received 11 August 2011; accepted 9 February 2012.

Published online 18 March 2012.

- Demetri, G. D. *et al.* Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *N. Engl. J. Med.* **347**, 472–480 (2002).
- Huang, M. E. *et al.* Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood* **72**, 567–572 (1988).
- Maemondo, M. *et al.* Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N. Engl. J. Med.* **362**, 2380–2388 (2010).
- Mok, T. S. *et al.* Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957 (2009).
- Ahrendt, S. A. *et al.* p53 mutations and survival in stage I non-small-cell lung cancer: results of a prospective study. *J. Natl. Cancer Inst.* **95**, 961–970 (2003).
- Gill, R. K. *et al.* Frequent homozygous deletion of the *LKB1/STK11* gene in non-small cell lung cancer. *Oncogene* **30**, 3784–3791 (2011).
- Ji, H. *et al.* LKB1 modulates lung cancer differentiation and metastasis. *Nature* **448**, 807–810 (2007).
- Nagata, Y. *et al.* PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* **6**, 117–127 (2004).
- Steels, E. *et al.* Role of p53 as a prognostic factor for survival in lung cancer: a systematic review of the literature with a meta-analysis. *Eur. Respir. J.* **18**, 705–719 (2001).

- Yeh, T. C. *et al.* Biological characterization of ARRY-142886 (AZD6244), a potent, highly selective mitogen-activated protein kinase kinase 1/2 inhibitor. *Clin. Cancer Res.* **13**, 1576–1583 (2007).
- Matsumoto, S. *et al.* Prevalence and specificity of *LKB1* genetic alterations in lung cancers. *Oncogene* **26**, 5911–5918 (2007).
- Weir, B. A. *et al.* Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450**, 893–898 (2007).
- Douillard, J. Y. *et al.* Molecular predictors of outcome with gefitinib and docetaxel in previously treated non-small-cell lung cancer: data from the randomized phase III INTEREST trial. *J. Clin. Oncol.* **28**, 744–752 (2010).
- Mascaux, C. *et al.* The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br. J. Cancer* **92**, 131–139 (2005).
- Ji, H. *et al.* Mutations in BRAF and KRAS converge on activation of the mitogen-activated protein kinase pathway in lung cancer mouse models. *Cancer Res.* **67**, 4933–4939 (2007).
- Verhoeven, D., Teijaro, J. R. & Farber, D. L. Pulse-oximetry accurately predicts lung pathology and the immune response during influenza infection. *Virology* **390**, 151–156 (2009).
- Li, D. *et al.* Bronchial and peripheral murine lung carcinomas induced by T790M-L858R mutant EGFR respond to HKI-272 and rapamycin combination therapy. *Cancer Cell* **12**, 81–93 (2007).
- Dykes, D. J., Bissery, M. C., Harrison, S. D. Jr & Waud, W. R. Response of human tumor xenografts in athymic nude mice to docetaxel (RP 56976, Taxotere). *Invest. New Drugs* **13**, 1–11 (1995).
- Engelman, J. A. *et al.* Effective use of PI3K and MEK inhibitors to treat mutant *Kras* G12D and *PIK3CA* H1047R murine lung cancers. *Nature Med.* **14**, 1351–1356 (2008).
- Berghmans, T. *et al.* Primary tumor standardized uptake value (SUV_{max}) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. *J. Thorac. Oncol.* **3**, 6–12 (2008).
- Vansteenkiste, J. F. *et al.* Prognostic importance of the standardized uptake value on ¹⁸F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: an analysis of 125 cases. Leuven Lung Cancer Group. *J. Clin. Oncol.* **17**, 3201–3206 (1999).
- Carretero, J. *et al.* Integrative genomic and proteomic analyses identify targets for *Lkb1*-deficient metastatic lung tumors. *Cancer Cell* **17**, 547–559 (2010).
- Hainsworth, J. D. *et al.* A phase II, open-label, randomized study to assess the efficacy and safety of AZD6244 (ARRY-142886) versus pemetrexed in patients with non-small cell lung cancer who have failed one or two prior chemotherapeutic regimens. *J. Thorac. Oncol.* **5**, 1630–1636 (2010).
- Erasmus, J. J., Rohren, E. & Swisher, S. G. Prognosis and reevaluation of lung cancer by positron emission tomography imaging. *Proc. Am. Thorac. Soc.* **6**, 171–179 (2009).
- Singh, M. *et al.* Assessing therapeutic responses in *Kras* mutant cancers using genetically engineered mouse models. *Nature Biotechnol.* **28**, 585–593 (2010).
- Tuveson, D. & Hanahan, D. Translational medicine: cancer lessons from mice to humans. *Nature* **471**, 316–317 (2011).
- Politi, K. & Pao, W. How genetically engineered mouse tumor models provide insights into human cancers. *J. Clin. Oncol.* **29**, 2273–2281 (2011).
- Nishino, M. *et al.* CT tumor volume measurement in advanced non-small-cell lung cancer: Performance characteristics of an emerging clinical tool. *Acad. Radiol.* **18**, 54–62 (2011).
- Zhao, B. *et al.* A pilot study of volume measurement as a method of tumor response evaluation to aid biomarker development. *Clin. Cancer Res.* **16**, 4647–4653 (2010).
- LoRusso, P. M. *et al.* Phase I and pharmacokinetic study of lapatinib and docetaxel in patients with advanced cancer. *J. Clin. Oncol.* **26**, 3051–3056 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is supported by the National Institutes of Health (CA122794, CA140594, CA137181, CA137008, CA147940, CA137008-01, 1U01CA141576, Lung SPOR P50CA090578), United against Lung Cancer Foundation, American Lung Association and Susan Spooner Research Fund.

Author Contributions Z.C., K.C., Z.W., Y.W., H.E., T.S., Y.L., T.T., J.O., J.L., P.G., M.S.W., C.X., M.Y., A.A., S.W., C.L., Y.N., C.G.P., Y.S., Y.F., C.Y., A.S., M.D.C., D.N.H., M.D.W., P.J.R., C.B.L., N.B., N.E.S., D.H.C., G.D.D., P.A.J., L.C.C., C.B.L., M.N. and P.P.P. performed experimental work and data analyses. M.B., L.R.C., D.B.C. and D.J. collected data and provided patient materials. A.L.K., J.A.E. and K.-K.W. conceived and supervised all aspects of the project. All authors contributed to the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.-K.W. (kwong1@partners.org), J.A.E. (jengelma@partners.org) or A.L.K. (andrew_kung@dfci.harvard.edu).

The mechanism of OTUB1-mediated inhibition of ubiquitination

Reuven Wiener¹, Xiangbin Zhang¹, Tao Wang^{1†} & Cynthia Wolberger¹

Histones are ubiquitinated in response to DNA double-strand breaks (DSB), promoting recruitment of repair proteins to chromatin¹. UBC13 (also known as UBE2N) is a ubiquitin-conjugating enzyme (E2) that heterodimerizes with UEV1A² (also known as UBE2V1) and synthesizes K63-linked polyubiquitin (K63Ub) chains at DSB sites in concert with the ubiquitin ligase (E3), RNF168 (ref. 3). K63Ub synthesis is regulated in a non-canonical manner by the deubiquitinating enzyme, OTUB1 (OTU domain-containing ubiquitin aldehyde-binding protein 1), which binds preferentially to the UBC13~Ub thiolester⁴. Residues amino-terminal to the OTU domain, which had been implicated in ubiquitin binding⁴, are required for binding to UBC13~Ub and inhibition of K63Ub synthesis⁵. Here we describe structural and biochemical studies elucidating how OTUB1 inhibits UBC13 and other E2 enzymes. We unexpectedly find that OTUB1 binding to UBC13~Ub is allosterically regulated by free ubiquitin, which binds to a second site in OTUB1 and increases its affinity for UBC13~Ub, while at the same time disrupting interactions with UEV1A in a manner that depends on the OTUB1 N terminus. Crystal structures of an OTUB1-UBC13 complex and of OTUB1 bound to ubiquitin aldehyde and a chemical UBC13~Ub conjugate show that binding of free ubiquitin to OTUB1 triggers conformational changes in the OTU domain and formation of a ubiquitin-binding helix in the N terminus, thus promoting binding of the conjugated donor ubiquitin in UBC13~Ub to OTUB1. The donor ubiquitin thus cannot interact with the E2 enzyme, which has been shown to be important for ubiquitin transfer^{6,7}. The N-terminal helix of OTUB1 is positioned to interfere with UEV1A binding to UBC13, as well as with attack on the thiolester by an acceptor ubiquitin, thereby inhibiting K63Ub synthesis. OTUB1 binding also occludes the RING E3 binding site on UBC13, thus providing a further component of inhibition. The general features of the inhibition mechanism explain how OTUB1 inhibits other E2 enzymes⁴ in a non-catalytic manner.

OTUB1 was previously identified as a K48 linkage-specific deubiquitinating enzyme that contains two distinct ubiquitin-binding sites (Fig. 1a): a distal site and a proximal site that includes the ~45 N-terminal residues of OTUB1 (ref. 5). These residues are important for OTUB1 inhibition of E2 activity⁴ and are absent in OTUB2, which does not inhibit UBC13 (ref. 4). It was previously shown that binding of the covalent inhibitor, ubiquitin aldehyde (Ubal), to the distal ubiquitin-binding site of OTUB1 stimulates binding of ubiquitin vinyl sulfone to the N terminus⁵. Because the OTUB1 N terminus was implicated in binding to the donor ubiquitin in the UBC13~Ub conjugate⁴, we asked whether Ubal binding to OTUB1 could enhance inhibition of UBC13 by stimulating binding of the OTUB1 N terminus to the donor ubiquitin. The results (Fig. 1b) showed a marked enhancement of the ability of OTUB1 to suppress K63Ub synthesis, indicating that Ubal is an allosteric effector that increases the affinity of the OTUB1 N terminus for the ubiquitin in the UBC13~Ub thiolester. This prompted us to ask whether free ubiquitin binding to the OTUB1 distal site could similarly stimulate binding of OTUB1 to UBC13~Ub

conjugates. To test this, we generated a mixture of charged and uncharged UBC13(C87S), which forms a more stable UBC13~Ub oxyester, purified away the free ubiquitin, and performed pull-down assays with H₆-OTUB1 in the presence and absence of added free ubiquitin. Remarkably, OTUB1 shows no preference for the charged UBC13~Ub in the absence of ubiquitin, whereas addition of 100 μ M free ubiquitin greatly enhances OTUB1 binding to UBC13~Ub, but not to uncharged UBC13 (Fig. 1c). By contrast, ubiquitin bearing hydrophobic patch mutations I44A, L8A or L8A/I44A/R42A (but not R42A alone) do not stimulate OTUB1 binding to UBC13~Ub like wild-type ubiquitin (Fig. 1c). The relative binding of OTUB1 to UBC13~Ub increases as the concentration of free ubiquitin is increased from 2 to 50 μ M (Supplementary Fig. 2). To verify that ubiquitin binding to the distal site of OTUB1 is important for inhibition of UBC13, we assayed the effect of distal site mutations, which were chosen based on structures of a covalent yeast Otu1-ubiquitin complex⁸ and of human OTUB1 (ref. 9). Distal site substitutions F193W, F193R and H217W disrupted the ability of OTUB1 to inhibit polyubiquitination by UBC13~UEV1A (Fig. 1d) without affecting binding of OTUB1 to UBC13 (Supplementary Fig. 3). Taken together, our results indicate that the ability of OTUB1 to bind preferentially to the UBC13~Ub conjugate and inhibit ubiquitin transfer is allosterically regulated by free ubiquitin binding to the distal site of OTUB1 (Fig. 1a), which triggers capture of the conjugated ubiquitin in the OTUB1 proximal site.

Because ubiquitin aldehyde most probably enhances interactions between the OTUB1 N terminus and the donor ubiquitin in UBC13~Ub, we examined the effect of N-terminal deletions in OTUB1 to delimit the minimal fragment needed for binding and inhibition. Deletion of residues 1–15 has no effect on inhibition of K63Ub synthesis by UBC13~UEV1A (Fig. 1e) whereas deletion of 30, 37 or 41 residues significantly disrupts inhibition. The OTUB1 Δ 15 deletion similarly behaves like full-length OTUB1 in pull-downs with the UBC13~Ub conjugate whereas larger deletions exhibit defects (Supplementary Fig. 4), indicating that N-terminal residues 16–45 are sufficient for activity.

Because a UEV (ubiquitin E2 variant) must bind to UBC13 and position the acceptor ubiquitin for K63Ub synthesis to occur¹⁰, we asked whether OTUB1 could bind to UBC13 in the presence of UEV1A. In gel filtration assays using fluorescently labelled UEV1A, OTUB1 and uncharged UBC13 migrate as a ternary complex with UEV1A (Fig. 1f). To assay binding to charged UBC13, we generated a non-hydrolysable conjugate in which Ub with a carboxy-terminal G75C is covalently linked to the active-site cysteine of UBC13 with dichloroacetone (DCA)¹¹. UEV1A binds to UBC13^{DCA}~Ub but OTUB1-Ubal interferes with UEV1A binding to the UBC13^{DCA}~Ub conjugate (Fig. 1g). By contrast, the N-terminal deletion, OTUB1 Δ 37, can still form a complex with UBC13^{DCA}~Ub and labelled UEV1 in the presence of Ubal (Fig. 1h), indicating that the N terminus of OTUB1 competes with UEV binding when OTUB1 is bound to Ubal. We verified that free ubiquitin has a similar effect on UEV binding by

¹Department of Biophysics and Biophysical Chemistry and the Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [†]Present address: National Institute of Allergy and Infectious Diseases, Vaccine Research Center, 40 Convent Drive, Bethesda, Maryland 20892, USA.

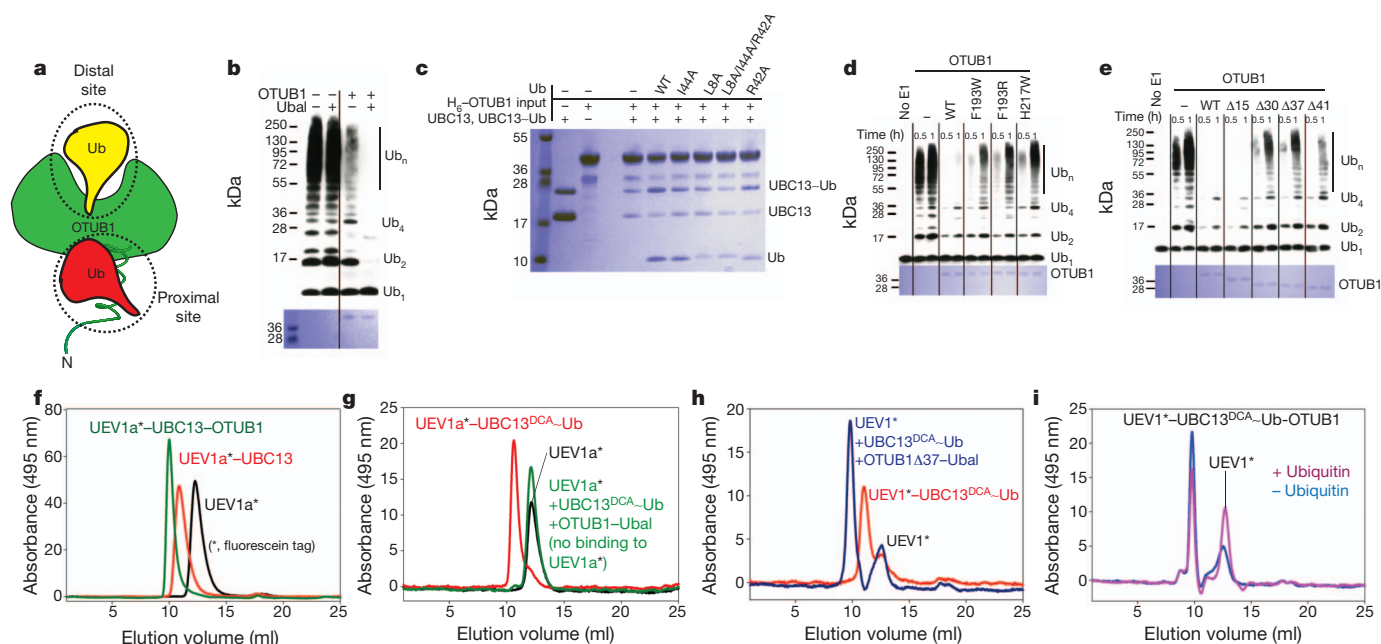


Figure 1 | Allosteric regulation of OTUB1 by ubiquitin. **a**, Schematic diagram of OTUB1 illustrating proximal and distal ubiquitin binding sites. **b**, Effect of ubiquitin aldehyde (Ubal) on the ability of human OTUB1 to inhibit K63 polyubiquitin synthesis by UBC13–UEV1A. Assays include 0.1 μ M E1, 0.4 μ M UBC13–UEV1A, 0.5 μ M human OTUB1, 5 μ M ubiquitin. The 3 h time point is shown in the presence (right) and absence (left) of human OTUB1, without (–) and with (+) 0.5 μ M Ubal. Top shows detection by anti-Ub western blot; Coomassie staining below shows level of human OTUB1. **c**, Pull-down assay showing binding of H_6 -tagged human OTUB1 to a mixture of UBC13 and UBC13~Ub oxyster in the presence and absence of 100 μ M free ubiquitin (wild type (WT) or mutant). **d**, Effect of human OTUB1 distal site mutations on inhibition of K63Ub synthesis. Assay performed as in **b** but with

1 μ M OTUB1. **e**, Effect of human OTUB1 N-terminal deletions of 15, 30, 37 and 42 residues on inhibition of K63Ub synthesis by UBC13–UEV1A. Assay performed as in **d**. **f**, Gel filtration showing complex formation between fluorescein-labelled UEV1A (UEV1A*), UBC13 and human OTUB1. Signal due to UEV1A only was monitored at 495 nm. **g**, Experiment performed as in **f** showing binding to UEV1A* by UBC13^{DCA}~Ub(G75C) in the absence (red) and presence (green) of OTUB1–Ubal. **h**, Experiment performed as in **f** but with human OTUB1 Δ 37. The position at which free UEV1* migrates is indicated. **i**, Experiment performed as in **f** with fluorescein-labelled UEV mixed with UBC13^{DCA}~Ub(G75C) and OTUB1 samples prepared in the presence and absence of 200 μ M ubiquitin. The position at which free UEV1* migrates is indicated.

comparing migration of a sample containing labelled UEV1, UBC13^{DCA}~Ub and OTUB1 prepared in the presence and absence of free ubiquitin and found that the ratio of free UEV1 to UEV1–UBC13^{DCA}~Ub–OTUB1 increases when ubiquitin is present (Fig. 1i). Similarly, pull-downs with H_6 –OTUB1 do not show an enhancement in coprecipitation of UEV1A along with UBC13~Ub in the presence of added free ubiquitin (Supplementary Fig. 5). These results indicate that the N terminus of OTUB1 interferes with UEV binding and thus with K63Ub synthesis, and that the ability of the N terminus to interfere with UEV depends upon a conformational change that is triggered by binding of free ubiquitin to OTUB1.

To determine the structural basis for OTUB1 inhibition of E2 enzymes, and how ubiquitin allosterically regulates OTUB1 activity, we determined the structure of *Caenorhabditis elegans* OTUB1 (worm OTUB1) bound to human UBC13 at 1.8 Å resolution (Fig. 2a), and a 2.35 Å resolution quaternary complex structure containing worm OTUB1, Ubal and a UBC13^{DCA}~Ub conjugate generated with Ub(G76C). The resulting non-native linkage is four bond lengths longer than the native thioester (Supplementary Fig. 6). Human UBC13 is 89% identical to worm UBC13, whereas human OTUB1 shares 34% sequence identity and 56% similarity with worm OTUB1 (Supplementary Fig. 7) and inhibits K63Ub chain formation by human UBC13–UEV1A (Supplementary Fig. 8a). Worm OTUB1 is a weaker inhibitor of UBC13, as reflected in its higher K_d of 58.5 μ M compared to 7.04 μ M for human OTUB1 (Supplementary Fig. 8b). Crystals of the worm OTUB1–Ubal–UBC13^{DCA}~Ub complex contain four complexes in the P2₁2₁2₁ asymmetric unit. The ubiquitin conjugated to UBC13 could be unambiguously positioned in two of the four complexes (Supplementary Fig. 9); our discussion focuses on the complex with the most well-ordered ubiquitin (complex 1). Because the N terminus of OTUB1 that plays a key role in inhibition is poorly conserved between

human and worm OTUB1, we also determined the 3.1 Å resolution structure of a quaternary complex with a hybrid OTUB1 containing the N-terminal 45 residues of human OTUB1 and the OTU domain of worm OTUB1 (Supplementary Fig. 7b). The hybrid human/worm OTUB1 inhibits K63Ub synthesis by UBC13–UEV1A (Supplementary Fig. 10). Details on all structure determinations are in Supplementary Methods and statistics are in Supplementary Table 1.

In the structure of apo worm OTUB1 bound to UBC13 (Fig. 2a), the OTU domain of worm OTUB1 binds to UBC13 in an orientation that places their respective active-site cysteines 28 Å apart on the same face of the complex, burying 1,280 Å² of total surface area. Of the 12 worm OTUB1 side chains at the interface with UBC13 (Fig. 2b), seven are identical in human OTUB1 and four are similar (Supplementary Fig. 7a) and can mediate comparable interactions with UBC13. Consistent with this, the double substitution Y170A/F138A in human OTUB1 (Y168A/F135A in worm OTUB1) is defective in binding to UBC13 (Supplementary Fig. 11). Similar interactions could form between OTUB1 and UBE2D2 (also known as UBCH5B) (Fig. 2c), but clashes due to an insertion and a non-conserved lysine would arise with UBE2L3 (also known as UBCH7), consistent with the observation that OTUB1 inhibits UBCH5B but not UBCH7 (ref. 4).

An overview of the human/worm OTUB1–Ubal–UBC13^{DCA}~Ub complex is shown in Fig. 2d, e. Ubal binds to the OTUB1 distal site while the donor ubiquitin in the UBC13~Ub conjugate binds in the OTUB1 proximal site, which comprises residues in both the OTU domain and the N terminus. In the absence of bound ubiquitin, the worm OTUB1 N terminus (residues 1–37, corresponding to human OTUB1 residues 1–39) is disordered (Fig. 2a). However, in the OTUB1–Ubal–UBC13^{DCA}~Ub complexes, part of the N terminus of OTUB1 becomes ordered, forming a ubiquitin-binding helix that contacts the donor ubiquitin in the distal site (Fig. 2e). Additional

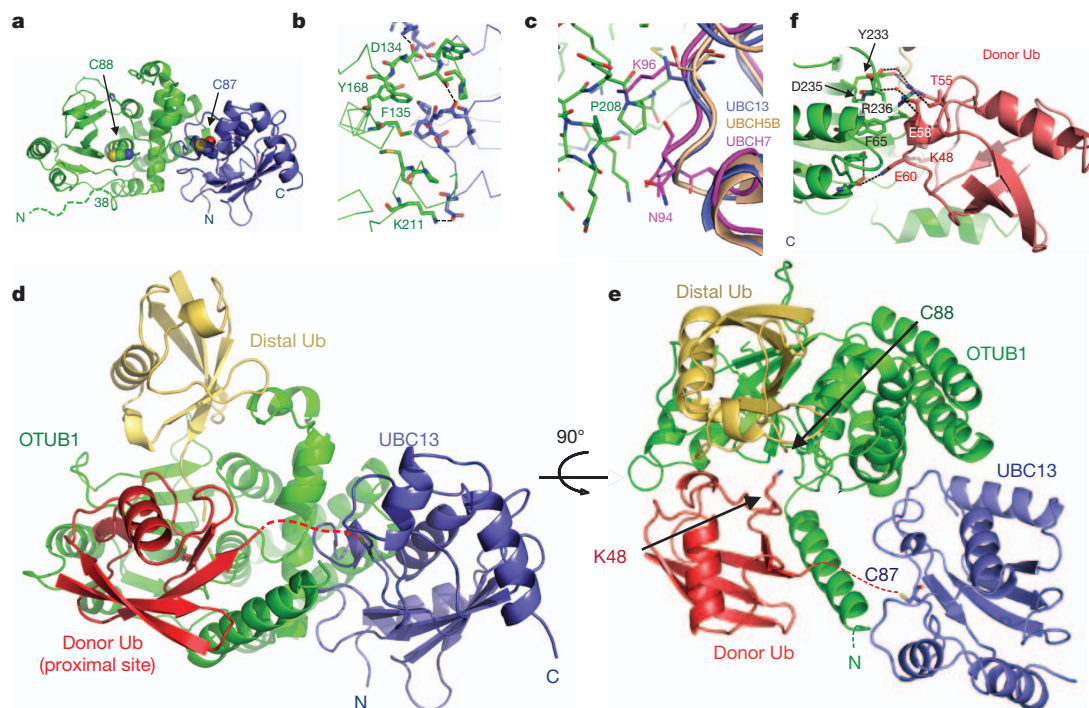


Figure 2 | Structure OTUB1–UBC13 and OTUB1–Uba1–UBC13^{DCA}~Ub. **a**, Complex of worm OTUB1 (green) bound to human UBC13 (blue). Respective active-site cysteines are shown as space-filling representations. Dashed line indicates disordered residues. **b**, Contacts at worm OTUB1 (green)–UBC13 (blue) interface. **c**, Superposition of UBCH5B (UBE2D2, PDB ID 2ESK) and UBCH7 (UBE2L3, PDB ID 1FBV) with UBC13 in the complex with worm OTUB1. UBCH7 contains an insertion (at N94) and a lysine (K96) that would interfere with binding. **d**, Structure of hybrid human/worm OTUB1

(green) bound to Uba1 (distal Ub, yellow), UBC13 (blue) and ubiquitin (proximal Ub, red) that is covalently linked to the active-site cysteine (C87) of UBC13 by a DCA linkage. Dashed line indicates disordered C-terminal residues 73–76 of the donor ubiquitin and DCA linkage. **e**, A 90° rotation compared to **d** showing positions of worm OTUB1 and UBC13 active-site cysteine and modelled location of K48 of the proximal ubiquitin. **f**, Contacts between the donor ubiquitin (red) and the OTU domain (green) in the worm OTUB1–Uba1–UBC13^{DCA}~Ub complex.

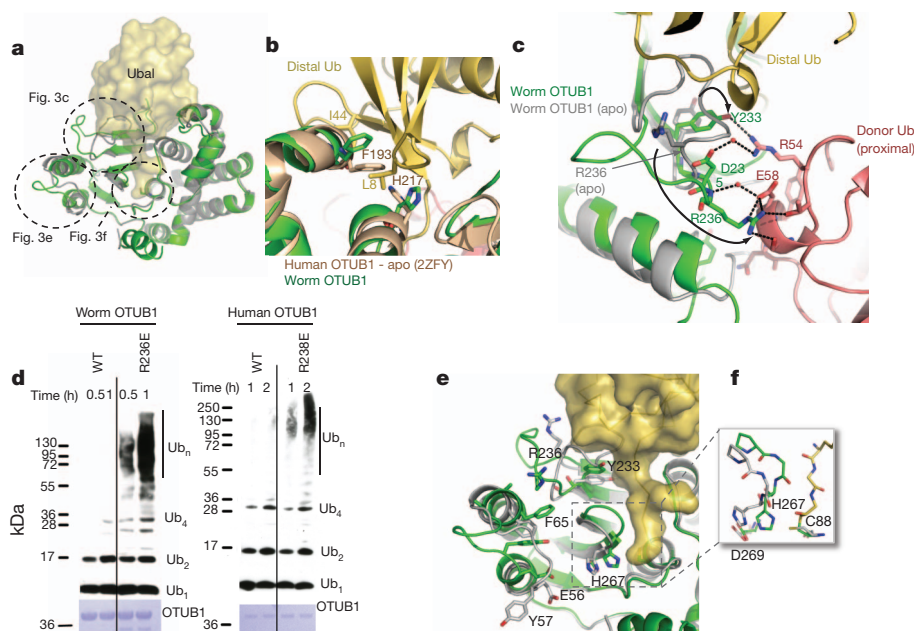


Figure 3 | Conformational changes in the OTU domain triggered by Uba1 binding. **a**, Superposition of worm OTUB1 (green) bound to Uba1 (yellow surface) with the structure of apo worm OTUB1 (grey). Dotted circles indicate regions of conformational change, which are illustrated in the figure panels noted. **b**, Location of human OTUB1 distal site mutations that affect inhibition. The structure of human OTUB1 (2ZFV; brown) is superimposed on worm OTUB1 (green)–Uba1 (yellow). Ubiquitin residues L8 and I44, where substitutions with alanine disrupt allosteric effect of ubiquitin binding, are shown. View is 180° rotation about vertical compared with **a**. **c**, Structural differences in the OTU domain in the presence (green) and absence (grey) of

distal Ub that affect contacts with the donor Ub. Arrows indicate conformational changes. Dotted lines indicate hydrogen bonds and salt bridges. View shown is from 'top' of complex as shown on right of panel **a**, rotated 90° counter-clockwise. **d**, Effect of mutating OTUB1 conserved arginine, worm OTUB1(R236E) and human OTUB1(R238E), on inhibition of UBC13–UEV1A. Assay performed as in Fig. 1b, with 1 μM human OTUB1 and 15 μM worm OTUB1. **e**, View of OTU domain structural rearrangements coloured as in **c**. View as in panel **a**; proximal ubiquitin not shown. **f**, Detailed view of catalytic triad in the presence and absence of Uba1 (carbon coloured as in **c**).

contacts with the donor ubiquitin are mediated by the OTU domain which, as described below, undergoes a set of conformational changes triggered by Ubal binding to the distal site.

The donor ubiquitin binds to the proximal site of OTUB1 (Fig. 2d) in an orientation that places K48 of the ubiquitin near the OTUB1 active site (Fig. 2e). A K48 isopeptide linkage can be modelled between the proximal and distal ubiquitins, consistent with OTUB1 isopeptidase specificity for K48-linked diubiquitin⁵. Residues 73–76 of the donor ubiquitin and the DCA linkage are not visible in the electron density map, indicating that they do not adopt a unique conformation in the crystal. The distance between the C-terminal ubiquitin residue and the active-site cysteine is approximately 12.5 Å, which is sufficient to accommodate the four missing residues and a native thiolester linkage. The donor ubiquitin interface with the OTU domain buries 850 Å² of surface area. Ubiquitin side chains that lie between residues 54–60 contact the OTU domain, forming both direct and water-mediated hydrogen bonds and van der Waals interactions (Fig. 2f). Three of the contacting worm OTUB1 side chains are R236, Y233 and D235, which are only in a position to contact ubiquitin in the quaternary complex.

The observed contacts between the donor ubiquitin and the OTU domain depend upon distal site binding of Ubal, which forms a covalent bond with the active-site cysteine (Supplementary Fig. 12) and triggers conformational changes in three regions of the globular OTU domain (Fig. 3a). Ubal binds to the distal ubiquitin binding site of OTUB1 (Fig. 3b) in a manner similar to yeast⁸ and viral^{12–14} OTU enzymes, and accounts for the effects of mutations in the OTUB1 distal site (Fig. 1d). A loop (residues 235–245) that partially occludes the distal site in the absence of ubiquitin undergoes a large rearrangement that relieves steric clashes with the distal ubiquitin and positions R236 of worm OTUB1 to make multiple contacts with the donor ubiquitin bound in the proximal site of OTUB1 (Fig. 3c). In the structure of apo human OTUB1 (ref. 9), this residue is disordered (backbone and side-chain atoms) and lies in a loop that presumably changes conformation upon distal ubiquitin binding. Mutating the conserved arginine to glutamic acid in both human (R238E) and worm (R236E) OTUB1 disrupts inhibition (Fig. 3d), consistent with its role in binding the donor ubiquitin. Interestingly, the corresponding residue is a glutamic acid in OTUB2, which lacks an N-terminal arm and does not inhibit UBC13 (ref. 4). Y233, which occludes the distal site in apo worm OTUB1 and undergoes a conformational change to hydrogen bond with the distal Ub (Fig. 3c), is conserved in human OTUB1 (Supplementary Fig. 7a). Another set of conformational changes in the loop connecting helices 1 and 2 of OTUB1 flips the solvent exposed Y57 side chain into the interior of the OTU domain, where it stacks between F65 and E56 (Fig. 3e). The altered loop conformation relieves steric clashes with the donor ubiquitin that would otherwise occur. Binding of the distal ubiquitin is accompanied by additional local rearrangements that narrow the binding cleft around the ubiquitin C-terminal tail (Fig. 3e) and moves the worm OTUB1 active-site histidine, H267, into a position between D269 and C88 to activate the cysteine for catalysis (Fig. 3f).

The OTUB1 N-terminal ubiquitin-binding helix seen in the structure spans residues 28–39 of worm OTUB1 (complex 1) and 25–44 of human OTUB1 (Figs 4a–c), burying 542 Å² and 626 Å², respectively, on the donor ubiquitin (electron density shown in Supplementary Fig. 13). The helix interacts with the donor ubiquitin in a manner reminiscent of the RAP80 UIM¹⁵ (Fig. 4d). Despite limited sequence identity between the worm OTUB1 and human OTUB1 N terminus (Fig. 4a), the three side chains that contact the donor ubiquitin in the 2.35 Å resolution structure of worm OTUB1 (Fig. 4b) are conserved in human OTUB1 (Fig. 4a) and are oriented towards ubiquitin in the same manner in the 3.1 Å resolution human/worm OTUB1 structure (Fig. 4c). In the worm OTUB1 complex (Fig. 4b), residues E37 and I34 contact donor ubiquitin residue H68 while Q33 interacts with backbone atoms. In the structure containing the human N terminus, the helix extends beyond the donor ubiquitin and approaches the UBC13 active-site cysteine (Fig. 4c). It is possible that additional residues may

be ordered when the complex is in solution, as nine residues from the minimal human OTUB1Δ15 fragment that exhibits full activity (Fig. 1d) are missing from the human/worm OTUB1 complex structure. It is not clear whether the shorter helix observed in the worm OTUB1 complex reflects a structural difference in solution, or whether crystal contacts interfere with helix formation. The close approach of the OTUB1 N terminus to the donor ubiquitin C terminus in both complexes (Figs 4b, c) leaves open the possibility that additional contacts may form with the donor ubiquitin tail linked to UBC13 via a native thiolester.

The structures show how OTUB1 interferes with UEV binding and positioning of the acceptor ubiquitin, and also occludes the RING E3

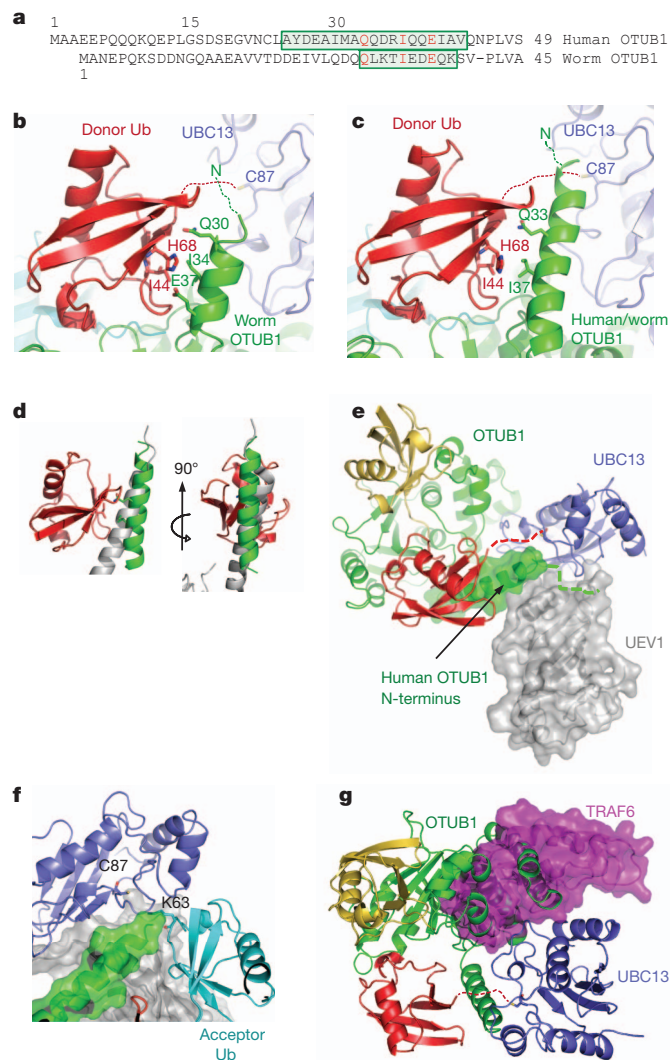


Figure 4 | OTUB1 N-terminal arm and the mechanism of E2 inhibition.

a, Sequence alignment of N-terminal arms of human OTUB1 and worm OTUB1. Boxed residues form a helix in the quaternary complex structures containing Ubal and UBC13^{DCA}~Ub; additional shaded residues in worm OTUB1 are ordered in complex 1 but are not helical. **b**, Donor Ub (red) interactions with the worm OTUB1 N-terminal helix (green); UBC13 shown in blue. Dashed lines indicate disordered residues. **c**, Interactions with the human OTUB1 N-terminal helix of the human/worm OTUB1 hybrid, depicted as in **b**. **d**, Superposition comparing RAP80 (grey, PDB ID 3A1Q) binding to ubiquitin (red) with human OTUB1 N-terminal helix (green). Two views are shown. **e**, Superposition of human/worm OTUB1–Ubal–UBC13^{DCA}~Ub with UBC13–UEV1 (1J7D) showing predicted position of UEV1 (grey). The solvent-accessible surface of the human N-terminal arm residues of OTUB1 is depicted. **f**, Modelled position of attacking K63 in acceptor Ub (cyan) based on yeast Ubc13~Ub–Mms2 (2GMI). **g**, Superposition with quaternary complex showing relative position of the TRAF6 E3 ligase (3HCT).

binding site. Figure 4e shows a superposition with the structure of a UBC13–UEV1 complex¹⁶ showing that the N-terminal helix of human OTUB1 clashes with the expected location of UEV1. Modelling of the predicted position of the acceptor ubiquitin based on the structure of yeast UBC13~Ub–Mms2 (ref. 17) shows the N terminus of OTUB1 in a position to interfere with attack by the acceptor ubiquitin lysine on the thiolester (Fig. 4f). Because OTUB1 also inhibits UBCH5B⁴, which does not function with a UEV, we propose that the OTUB1 N terminus may also interfere with acceptor ubiquitin binding for other E2s. The repositioning of the donor ubiquitin away from the E2 also probably contributes to inhibition, in light of evidence that the donor ubiquitin in the E2–Ub thiolester interacts specifically with the E2 (refs 18, 19) and that this is essential for ubiquitin transfer^{6,7}. In addition, superposition with the structure of UBC13 bound to TRAF6 (ref. 20) shows that the OTUB1 binding site overlaps with the E3 RING-binding site (Fig. 4g), indicating that competition between OTUB1 and RNF168 would further suppress UBC13 activity *in vivo*. Competition with E3 binding is likely to be particularly important for OTUB1 inhibition of UBCH5B which, unlike UBC13, is strictly dependent upon an E3 ligase for activity.

The ability of OTUB1 to serve as both an isopeptidase and an inhibitor of E2 enzyme activity arises from its ability to bind to selected E2s, while taking advantage of the allosteric communication between the proximal and distal ubiquitin binding sites of OTUB1 and the distinctive features of its N terminus. Given the high degree of coupling between the multiple binding interactions within the OTUB1–Ub–UBC13~Ub complex, the degree of inhibition *in vivo* will clearly depend upon the relative concentrations of OTUB1, E2~Ub thiolester, E3 and free ubiquitin in the cell. An interesting question is whether the dependence of OTUB1 repression on ubiquitin binding to the distal site is exploited to modulate OTUB1 activity in response to fluctuations in the concentration of free ubiquitin or of free chains, whose C-terminal subunits could similarly bind to the distal site of OTUB1. Our findings establish new directions for investigating how the allosteric regulation of OTUB1 may be exploited to regulate ubiquitination in the DNA damage response.

METHODS SUMMARY

Cloning, expression, protein purification and crystallization are described in Methods and in accompanying references⁵. The DCA linkage between the active-site cysteine of UBC13 and a C-terminal cysteine in Ub(G75C) or Ub(G76C) was generated by a modification of the published method¹¹. The hybrid human/worm OTUB1 protein contains residues 1–45 of human OTUB1 and residues 43–276 of worm OTUB1. Structures were determined by molecular replacement as described in Methods. Free ubiquitin chain synthesis was assayed by gel electrophoresis and products were detected by western blot with anti-ubiquitin antibody or by Coomassie staining. Pull-down assays were performed with purified recombinant protein. Assays of complex formation between OTUB1, UBC13, UBC13^{DCA}~Ub and UEV1A were performed by gel filtration with fluorescein-labelled UEV1A or UEV1, monitoring fluorescein absorbance at 495 nm. Binding of OTUB1 to UBC13 was measured by fluorescence anisotropy using fluorescently labelled UBC13, and equilibrium dissociation constants were calculated using SigmaPlot (SPSS).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 November 2011; accepted 3 February 2012.

Published online 22 February 2012.

1. Al-Hakim, A. *et al.* The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair (Amst.)* **9**, 1229–1240 (2010).

2. Deng, L. *et al.* Activation of the IκB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103**, 351–361 (2000).
3. Stewart, G. S. *et al.* The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell* **136**, 420–434 (2009).
4. Nakada, S. *et al.* Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature* **466**, 941–946 (2010).
5. Wang, T. *et al.* Evidence for bidentate substrate binding as the basis for the K48 linkage specificity of otubain 1. *J. Mol. Biol.* **386**, 1011–1023 (2009).
6. Saha, A., Lewis, S., Kleiger, G., Kuhlman, B. & Deshaies, R. J. Essential role for ubiquitin-ubiquitin-conjugating enzyme interaction in ubiquitin discharge from Cdc34 to substrate. *Mol. Cell* **42**, 75–83 (2011).
7. Wickliffe, K. E., Lorenz, S., Wemmer, D. E., Kuriyan, J. & Rape, M. The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **144**, 769–781 (2011).
8. Messick, T. E. *et al.* Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. *J. Biol. Chem.* **283**, 11038–11049 (2008).
9. Edelmann, M. J. *et al.* Structural basis and specificity of human otubain 1-mediated deubiquitination. *Biochem. J.* **418**, 379–390 (2009).
10. Hofmann, R. M. & Pickart, C. M. Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645–653 (1999).
11. Yin, L., Krantz, B., Russell, N. S., Deshpande, S. & Wilkinson, K. D. Nonhydrolyzable diubiquitin analogues are inhibitors of ubiquitin conjugation and deconjugation. *Biochemistry* **39**, 10001–10010 (2000).
12. James, T. W. *et al.* Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease. *Proc. Natl Acad. Sci. USA* **108**, 2222–2227 (2011).
13. Capodagli, G. C. *et al.* Structural analysis of a viral ovarian tumor domain protease from the Crimean-Congo hemorrhagic fever virus in complex with covalently bonded ubiquitin. *J. Virol.* **85**, 3621–3630 (2011).
14. Akutsu, M., Ye, Y., Virdee, S., Chin, J. W. & Komander, D. Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains. *Proc. Natl Acad. Sci. USA* **108**, 2228–2233 (2011).
15. Sato, Y. *et al.* Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. *EMBO J.* **28**, 2461–2468 (2009).
16. Moraes, T. F. *et al.* Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2–hUbc13. *Nature Struct. Biol.* **8**, 669–673 (2001).
17. Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. & Wolberger, C. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Struct. Mol. Biol.* **13**, 915–920 (2006).
18. Wenzel, D. M., Stoll, K. E. & Klever, R. E. E2s: structurally economical and functionally replete. *Biochem. J.* **433**, 31–42 (2011).
19. McKenna, S. *et al.* An NMR-based model of the ubiquitin-bound human ubiquitin conjugation complex Mms2–Ubc13. The structural basis for lysine 63 chain catalysis. *J. Biol. Chem.* **278**, 13151–13158 (2003).
20. Yin, Q. *et al.* E2 interaction and dimerization in the crystal structure of TRAF6. *Nature Struct. Mol. Biol.* **16**, 658–666 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Henderson for generating the human/worm OTUB1 clone and C. Berndsen, A. DiBello, A. Datta and M. Bianchet for discussions. GM/CA-CAT has been funded in whole or in part with funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. DE-AC02-06CH11357.

Author Contributions R.W. and C.W. designed the experiments and R.W. performed all biochemical experiments. Cloning, expression and protein purification were carried out by X.Z., T.W. and R.W. Complexes were prepared for crystallization and crystals were grown by X.Z. and R.W.; R.W. determined the crystal structure with guidance from C.W.; R.W. and C.W. wrote the manuscript.

Author Information Coordinates and diffraction amplitudes are deposited in the Protein Data Bank under accession numbers 4DHI (worm OTUB1–UBC13), 4DHJ (worm OTUB1–Uba1–UBC13^{DCA}–Ub) and 4DHZ (human/worm OTUB1–Uba1–UBC13^{DCA}–Ub). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.W. (cwollberg@jhmi.edu).

METHODS

Cloning and mutagenesis. Cloning of human and *C. elegans* OTUB1 (human OTUB1 and worm OTUB1, respectively) was performed as described previously⁵. The human UBC13 open reading frame was amplified from a human complementary DNA library and cloned into a pET vector containing an N-terminal His₆-SUMO-2 tag (pETSUMO-2). The human UEV1A ORF was synthesized (Integrated DNA Technologies) and subcloned into the pETSUMO-2 vector as above. The human UEV1 (missing the first 30 residues of UEV1A) expression plasmid was purchased from Addgene (<http://www.addgene.org>).

Mutants of human OTUB1 were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene) following the manufacturer's protocol. The hybrid human/worm OTUB1 was generated by swapping the first 41 residues of worm OTUB1 with the first 45 residues of human OTUB1 using Infusion ligation-free cloning (Clontech). Human OTUB1 with an N-terminal 41-residue truncation (OTUB1ΔN41) was generated as previously described⁵, all other OTUB1 deletions were generated using Infusion ligation-free cloning (Clontech).

Protein expression and purification. All proteins were expressed in *Escherichia coli* Rosetta-2 (DE3) cells grown in Luria-Bertani (LB) medium. Cultures were inoculated using 1% (v/v) overnight saturated cultures and were grown at 37 °C to an OD₆₀₀ of 0.8. Proteins were induced at 16 °C overnight by addition of 1 mM isopropyl-β-D-thio-galactoside (IPTG). Cells were harvested by centrifugation (8,000g, 15 min) and either lysed immediately or stored at -80 °C for later use.

Human OTUB1, worm OTUB1, human E1 enzyme and ubiquitin were purified as previously described^{5,21}. Deletions and mutants of human and *C. elegans* OTUB1 and of ubiquitin were purified according to the same protocol as the wild-type proteins. UBC13 and UEV1A were purified by resuspending cell pellets in lysis buffer (20 mM HEPES pH 7.3, 300 mM NaCl, 10 mM imidazole, 2 mM β-mercaptoethanol) after adding 0.1 mM phenyl-methyl sulphonyl fluoride (PMSF). Cells were disrupted using a Microfluidizer (Microfluidics) and the lysate was centrifuged to remove cell debris. The lysate was subjected to immobilized metal affinity chromatography (IMAC) using 5 ml His-Trap columns (GE Biosciences) developed with a linear imidazole gradient of 25–400 mM in 20 column volumes. Fractions containing purified protein were pooled, SENP-2 protease was added in a ratio of 1:100 to cleave off the His-SUMO-2 tag, and pooled fractions were dialysed overnight at 4 °C against lysis buffer. Cleaved protein was then subjected to a second round of IMAC and the cleaved protein was collected from the flow-through. Proteins were then purified by gel filtration on a Superdex 75 column (GE Healthcare), dialysed into 20 mM HEPES, pH 7.3, 150 mM NaCl and 1 mM dithiothreitol (DTT), concentrated and stored at -80 °C. Proteins for crystallization, enzyme assays and binding studies were >98% pure as visualized on a Coomassie-stained gel. His₆-tagged human OTUB1 used in pull-down assays was ~90% pure.

Protein modifications. UBC13, UEV1A and UEV1 were labelled with fluorescein-5-maleimide (Invitrogen) as described in the manufacturer's protocol. Ubiquitin aldehyde was prepared as described²².

Preparation of UBC13~Ub conjugates. UBC13(C87S)~Ub oxyster was prepared as previously described¹⁷. The UBC13^{DCA}~Ub covalent conjugate was prepared according to a modification of the protocol from ref. 11. Purified ubiquitin containing the substitution G76C (Ub(G76C)) or G75C (Ub(G75C)) and UBC13 were dialysed separately overnight into 20 mM sodium borate buffer, pH 8.0 and 2 mM TCEP (tris(2-carboxyethyl)phosphine), mixed in the proportion of 1 mM Ub(G76C) or Ub(G75C) to 330 μM UBC13, and incubated on ice for 15 min. A stock of 20 mM 1,3-dichloroacetone (DCA) was prepared in dimethyl-formamide (DMF) and added to the conjugation reaction to a final concentration of 0.8 mM DCA. The reaction was stopped after 1 h by addition of 10 mM β-mercaptoethanol. The coupling efficiency was approximately 50%. For the Ub(G76C) reaction, the mix was diluted tenfold with 10 mM Tris, pH 8, loaded onto a mono Q column (GE Healthcare) pre-equilibrated with 10 mM Tris, pH 8. Free Ub(G76C) eluted in the flow-through and UBC13^{DCA}~Ub eluted together with unconjugated UBC13 in 180 mM NaCl in 20 mM Tris, pH 8. For the Ub(G75C) reaction, UBC13^{DCA}~Ub(G75C) was purified by gel filtration on a Superdex 75 column pre-equilibrated with 20 mM HEPES pH 7.3, 100 mM NaCl and 2 mM DTT. The separation efficiency was about 10% of the total amount of UBC13^{DCA}~Ub(G75C) in the reaction mix.

Purification of worm OTUB1-Ubal-UBC13^{DCA}~Ub(G76C) quaternary complex. Worm OTUB1 was incubated on ice with Ubal in a 1:4 molar ratio for 15 min and added to the purified apo human UBC13 and UBC13^{DCA}~Ub mixture such that UBC13^{DCA}~Ub was in twofold excess over worm OTUB1, as estimated by gel electrophoresis. The mixture was incubated for another 15 min on ice and loaded onto a Superdex 200 column (GE Healthcare) pre-equilibrated with 20 mM Tris, pH 7.45, 150 mM NaCl and 2 mM DTT. The OTUB1-Ubal-UBC13^{DCA}~Ub complex eluted as a single peak and was concentrated to 10 mg ml⁻¹ and stored at -80 °C.

Crystallization. All crystals were grown by the hanging-drop vapour diffusion method at 20 °C. A worm OTUB1-UBC13 complex was prepared by incubating worm OTUB1 and human UBC13 at a molar ratio of 1:1 and total protein concentration of 26 mg ml⁻¹ for 10 min at room temperature. Crystals were grown from a 1:1 mix of protein and well solution containing 100 mM sodium cacodylate, pH 6.5 and 1 M trisodium citrate and appeared in about 2–3 days. Crystals were transferred to cryoprotectant consisting of well solution plus 20% ethylene glycol and then flash-frozen in liquid nitrogen.

Crystals of the worm OTUB1-Ubal-UBC13^{DCA}~Ub complex were grown from a 1:1 mix of complex (10 mg ml⁻¹) with well solution containing 100 mM Bis-Tris pH 6.5, 23% PEG 3350 and 0.26–0.3 M sodium chloride. Crystals appeared in about 1–2 days, were cryoprotected by well solution with added 15% ethylene glycol and then flash-frozen in liquid nitrogen.

Crystals of the human/worm OTUB1-Ubal-UBC13^{DCA}~Ub complex were grown from a 1:1 mix of complex (10 mg ml⁻¹) with well solution containing 100 mM MES pH 6.5, 21% PEG 10,000 and 0.1 M sodium chloride. Crystals appeared in about 2–3 days, were cryoprotected by well solution with added 15% ethylene glycol and then flash-frozen in liquid nitrogen.

Structure determination. Diffraction data were recorded at the GM/CA-CAT beamline 23-ID-D/B at the Advanced Photon Source under standard cryogenic conditions and processed with iMOSFLM²³ for worm OTUB1-human UBC13 crystals and HKL2000²⁴ for the worm OTUB1-Ubal-UBC13^{DCA}~Ub crystals. For the worm OTUB1-Ubal-UBC13^{DCA}~Ub structure two data sets were collected from a single crystal, merged and processed with HKL2000²⁴. All data were collected at a wavelength of 1.033 Å. The structure of worm OTUB1-UBC13 was determined by molecular replacement with Phaser²⁵ using structures of UBC13 (1J7D) and apo human OTUB1 (2ZFY). The structure of worm OTUB1-Ubal-UBC13^{DCA}~Ub was determined by molecular replacement with Molrep²⁶ using structures of the worm OTUB1-human UBC13 complex and ubiquitin (from 2GMI). The initial molecular replacement search performed with the worm OTUB1-UBC13 complex located four complexes in the asymmetric unit. The resulting positions of the OTUB1-UBC13 complexes were then fixed and ubiquitin was used as search model to locate the four ubiquitin aldehydes in the crystal. The position of the worm OTUB1-Ubal-UBC13 complex was fixed and another search with ubiquitin (1-71) located two molecules of donor ubiquitin in the asymmetric unit. The structure of human/worm OTUB1-Ubal-UBC13^{DCA}~Ub was determined by molecular replacement with Phaser using one complex of worm OTUB1-Ubal-human UBC13~Ub lacking the first 42 residues of worm OTUB1.

All structures were subjected to multiple rounds of manual correction and refinement using COOT²⁷ and Phenix²⁸. The final stages of refinement for the worm OTUB1-Ubal-UBC13^{DCA}~Ub complex and human/worm OTUB1-Ubal-UBC13^{DCA}~Ub ternary complex were done using REFMAC5²⁹. Simulated annealing omit maps were calculated with CNS³⁰ and used to verify selected portions of the model.

The final model of worm OTUB1-human UBC13 complex includes residues 38–275 of worm OTUB1 and 3–152 of human UBC13. The final model of worm OTUB1-Ubal-UBC13^{DCA}~Ub includes four complexes in the asymmetric unit: two containing all four proteins (worm OTUB1, Ubal, human UBC13 and Ub) and two lacking the donor ubiquitin conjugated to UBC13. There is no density in any of the complexes corresponding to the five C-terminal amino acids of ubiquitin or to the DCA linkage, which connects ubiquitin to the human UBC13 active-site cysteine. The number of worm OTUB1 N-terminal residues visible in the map differs among the four complexes as follows: complex 1, 28–275; complex 2, 31–275; complex 3, 36–276; complex 4, 38–276. The final model of the human/worm OTUB1-Ubal-UBC13^{DCA}~Ub complex includes residues 20–275 of human/worm OTUB1, 3–151 of human UBC13, 1–76 of Ubal and 1–72 of ubiquitin.

Protein-protein interaction surfaces were analysed using the PISA server at EBI (<http://www.pdbe.org/PISA>) and manually inspected using COOT and PYMOL (<http://www.pymol.org>). Figures were generated with PYMOL.

Fluorescence polarization binding assay. Fluorescein-labelled human UBC13 (20 nM) was incubated with increasing concentrations of human OTUB1 wild type or mutants in 20 mM Tris, pH 7.6, 150 mM NaCl and 10 mM β-mercaptoethanol. Polarization measurements were recorded at 25 °C with an ISS Chronos Fluorescence Lifetime Spectrometer at excitation and emission wavelengths of 492 and 520 nm, respectively. Binding data were analysed and K_d values were calculated by nonlinear regression in SigmaPlot (SPSS).

UEV binding assay. UEV1A and UEV1 were fluorescein-labelled; all other proteins are unlabelled. The experiment was performed with an analytical Superdex 75 column pre-equilibrated with 20 mM HEPES pH 7.3, 100 mM NaCl and 2 mM DTT. Absorbance was detected at 495 nm to monitor the presence of labelled UEV1A or UEV1. For each run, proteins were incubated for 20 min on ice before loading onto the column. The protein concentrations used in the different experiments were: Fig. 1e, UEV1A 20 μM, UBC13 40 μM and human OTUB1

100 μM ; Fig. 1f, UEV1A 10 μM , UBC13~Ub 10 μM , human OTUB1 50 μM and Ubal 50 μM ; Fig. 1g, UEV1 20 μM , UBC13~Ub 20 μM , human OTUB1($\Delta 37$) 100 μM and Ubal 100 μM ; Fig. 1h, UEV1 20 μM , UBC13~Ub 20 μM , human OTUB1 100 μM and ubiquitin 200 μM .

In vitro ubiquitination assay. Ubiquitination assays were performed in 25 mM Tris-HCl (pH 8.0) buffer containing 0.1 mM DTT, 1 mM ATP, 2.5 mM MgCl_2 , 5 mM creatine phosphate, 0.3 units ml^{-1} inorganic pyrophosphatase, and 0.3 units ml^{-1} creatine kinase. Proteins in the amounts of 0.4 μM UBC13, 0.4 μM UEV1A and 5 μM ubiquitin were mixed with human OTUB1 (1 μM) or worm OTUB1 (15 μM). Reactions were initiated by the addition of 0.1 μM E1 enzyme, incubated at 37 °C, and stopped at different time points by adding denaturing SDS-PAGE loading dye containing β -mercaptoethanol (BME). For Fig. 1b, 0.5 μM human OTUB1 was incubated with 0.5 μM Ubal for 15 min before addition to the reaction. Reaction products were separated on a 4–12% Bis-Tris NuPAGE (Invitrogen) gel and transferred to a polyvinylidene fluoride (PVDF) membrane. Membranes were denatured in a 6 M guanidine HCl, 20 mM Tris-HCl, pH 7.5, 1 mM PMSF, 5 mM β -mercaptoethanol solution for 30 min at 4 °C and then washed extensively in Tris-buffered saline and Tween 20 (TBST). Membrane were blocked overnight at 4 °C with 5% BSA in TBST and incubated for 1 h with ubiquitin antibody (P4D1 Santa Cruz) 1:1,000 at room temperature followed by anti-mouse horseradish peroxidase (HRP)-conjugated secondary antibody. OTUB1 was detected with Coomassie brilliant blue or SimplyBlue SafeStain (Invitrogen).

Pull-down assays. Ni^{2+} -NTA beads were equilibrated in buffer A (50 mM phosphate buffer pH 8.0, 100 mM NaCl, 5 mM β -mercaptoethanol and 10 mM Imidazole). 6 \times His-human OTUB1 (30 μg) was incubated with pre-equilibrated beads in 200 μl of buffer A for 30 min. Beads were washed with 400 μl buffer A and incubated with a mixture of human UBC13 and human UBC13(C87S)~Ub with

and without the indicated concentration of free ubiquitin (2–100 μM) in 200 μl buffer A for 1 h. Beads were washed with 400 μl buffer A for 10 min and eluted with 25 μl of buffer A plus 250 mM imidazole. Eluates were analysed by gel electrophoresis and staining with Coomassie blue or SimplyBlue SafeStain (Invitrogen). The pull-down in Supplementary Fig. 2 was performed as above except for the addition of 6 \times His-human OTUB1 (7 μg), human UBC13(C87S)~Ub (7 μg) and ubiquitin as indicated in the figure.

21. Berndsen, C. E. & Wolberger, C. A spectrophotometric assay for conjugation of ubiquitin and ubiquitin-like proteins. *Anal. Biochem.* **418**, 102–110 (2011).
22. Hu, M. *et al.* Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J.* **24**, 3747–3756 (2005).
23. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D* **67**, 271–281 (2011).
24. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
25. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
26. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010).
27. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
28. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
29. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
30. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

The *Shigella flexneri* effector OspI deamidates UBC13 to dampen the inflammatory response

Takahito Sanada¹, Minsoo Kim¹, Hitomi Mimuro², Masato Suzuki³, Michinaga Ogawa³, Akiho Oyama³, Hiroshi Ashida³, Taira Kobayashi³, Tomohiro Koyama⁴, Shinya Nagai⁴, Yuri Shibata⁵, Jin Gohda⁵, Jun-ichiro Inoue⁵, Tsunehiro Mizushima⁶ & Chihiro Sasakawa^{1,3,4}

Many bacterial pathogens can enter various host cells and then survive intracellularly, transiently evade humoral immunity, and further disseminate to other cells and tissues. When bacteria enter host cells and replicate intracellularly, the host cells sense the invading bacteria as damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs) by way of various pattern recognition receptors. As a result, the host cells induce alarm signals that activate the innate immune system¹. Therefore, bacteria must modulate host inflammatory signalling and dampen these alarm signals^{2–4}. How pathogens do this after invading epithelial cells remains unclear, however. Here we show that OspI, a *Shigella flexneri* effector encoded by *ORF169b* on the large plasmid and delivered by the type III secretion system, dampens acute inflammatory responses during bacterial invasion by suppressing the tumour-necrosis factor (TNF)-receptor-associated factor 6 (TRAF6)-mediated signalling pathway. OspI is a glutamine deamidase that selectively deamidates the glutamine residue at position 100 in UBC13 to a glutamic acid residue. Consequently, the E2 ubiquitin-conjugating activity required for TRAF6 activation is inhibited, allowing *S. flexneri* OspI to modulate the diacylglycerol-CBM (CARD-BCL10-MALT1) complex-TRAF6-nuclear-factor- κ B signalling pathway. We determined the 2.0 Å crystal structure of OspI, which contains a putative cysteine-histidine-aspartic acid catalytic triad. A mutational analysis showed this catalytic triad to be essential for the deamidation of UBC13. Our results suggest that *S. flexneri* inhibits acute inflammatory responses in the initial stage of infection by targeting the UBC13-TRAF6 complex.

The rupture of host cell membranes by invasive bacteria such as *S. flexneri* and *Listeria monocytogenes* is sensed as DAMPs and triggers acute inflammatory responses by activating various cell signals^{5,6}. Within the cytoplasm, *S. flexneri* multiplies and subsequently spreads to neighbouring cells and, during this dissemination, the bacteria release peptidoglycan, lipopolysaccharide and flagellin^{2,3,7}. These bacterial components are recognized as PAMPs by cytoplasmic pattern recognition receptors, such as NOD-like receptors (NLRs), which induce nuclear factor- κ B (NF- κ B)-mediated and mitogen-activated protein kinase (MAPK)-mediated inflammatory responses³. To counteract this host defence, bacteria have evolved mechanisms to modulate host inflammatory responses³, through delivering several effectors that modulate inflammatory signalling by way of the type III secretion system (T3SS)^{8,9}. To better understand these mechanisms, we searched for additional *S. flexneri* effectors that modulate acute inflammatory responses to bacterial invasion, and we found that OspI has a pivotal role (Supplementary Fig. 1).

To assess the role of OspI, we carried out a comprehensive microarray analysis of HeLa cells infected with each of three *S. flexneri* strains: YSH6000 (wild type), Δ ospI (ospI-deficient) or S325 (T3SS-deficient).

We found that the messenger RNA levels of chemokines (for example, interleukin-8 (IL-8), CC-chemokine ligand 20 (CCL20), CXC-chemokine ligand 2 (CXCL2) and CCL2) and cytokines (tumour-necrosis- α (TNF- α) and IL-6) were greatly increased in Δ ospI-infected cells at 60 min after infection (Supplementary Fig. 2a). This elevated chemokine and cytokine production was detected as early as 30 min after infection (Fig. 1a and Supplementary Fig. 2b). Increased phosphorylation of inhibitor of NF- κ B (I κ B α) was detected in Δ ospI-infected HeLa cells relative to YSH6000-infected HeLa cells as early as 10 min after infection (Fig. 1b). The nuclear translocation of NF- κ B (p65 subunit) was fourfold higher in cells infected with Δ ospI compared with YSH6000 at 20 min after infection (Fig. 1c and Supplementary Fig. 3), suggesting that OspI can dampen the acute inflammatory response.

We next tested whether OspI inhibits NF- κ B activation when *S. flexneri* infects cells, and we found that ectopic OspI expression further inhibited NF- κ B activation on *S. flexneri* infection (Fig. 1d). When *S. flexneri* infects epithelial cells, NOD1-RIP2 (nucleotide-binding oligomerization domain 1-receptor-interacting serine/threonine kinase 2)-dependent and NOD1-RIP2-independent NF- κ B signalling pathways are both activated. PAMPs stimulate the dependent pathway¹⁰, whereas DAMPs stimulate the independent pathway⁶. Therefore, we investigated the levels of NF- κ B activation in HeLa cells transiently expressing OspI (designated HeLa/OspI cells) or a mock control. When HeLa/OspI cells were stimulated with TNF- α , NOD1 or phorbol myristate acetate (PMA), an examination of the levels of activated NF- κ B revealed that OspI suppressed PMA-mediated NF- κ B activation but not TNF α - or NOD1-mediated NF- κ B activation (Fig. 1d).

Because PMA is a substitute for diacylglycerol (DAG) in the activation of the protein kinase C (PKC)-NF- κ B pathway¹¹, and because DAG in the host membrane can trigger antibacterial autophagy against *Salmonella enterica* serovar Typhimurium¹², we examined the membrane ruffles protruding around *S. flexneri* entry sites in cells expressing PKC-C1-GFP (the PKC-C1 region fused to green fluorescent protein (GFP) as a DAG sensor)¹³. We found that DAG accumulated around the bacterial entry site (Supplementary Fig. 4a). Indeed, the increased IL8 mRNA production in Δ ospI-infected cells was suppressed by treating the cells with propranolol, a DAG inhibitor (Supplementary Fig. 4b).

The DAG-NF- κ B pathway is mediated through the CBM complex, a major regulator of NF- κ B signalling in lymphoid, myeloid and non-myeloid cells in innate and adaptive immune responses¹⁴. We examined whether OspI modulates CBM-complex-mediated NF- κ B signalling, through knocking down BCL10 production with shorting interfering RNA (siRNA), and we found that the IL8 mRNA levels were greatly decreased compared with cells treated with control siRNA (Fig. 1e and Supplementary Fig. 4c). GFP-MALT1 was recruited to the *S. flexneri* entry point in HeLa/GFP-MALT1 cells because MALT1 functionally

¹Department of Infectious Disease Control, International Research Center for Infectious Diseases, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. ²Division of Bacteriology, Department of Infectious Disease Control, International Research Center for Infectious Diseases, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. ³Division of Bacterial Infection, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. ⁴Nippon Institute for Biological Science, 9-2221-1 Shinmachi, Ome, Tokyo 198-0024, Japan. ⁵Division of Cellular and Molecular Biology, Department of Cancer Biology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan. ⁶Picobiology Institute, Department of Life Science, Graduate School of Life Science, University of Hyogo, 3-2-1 Kouto, Kamigori, Akoh, Hyogo 678-1297, Japan.

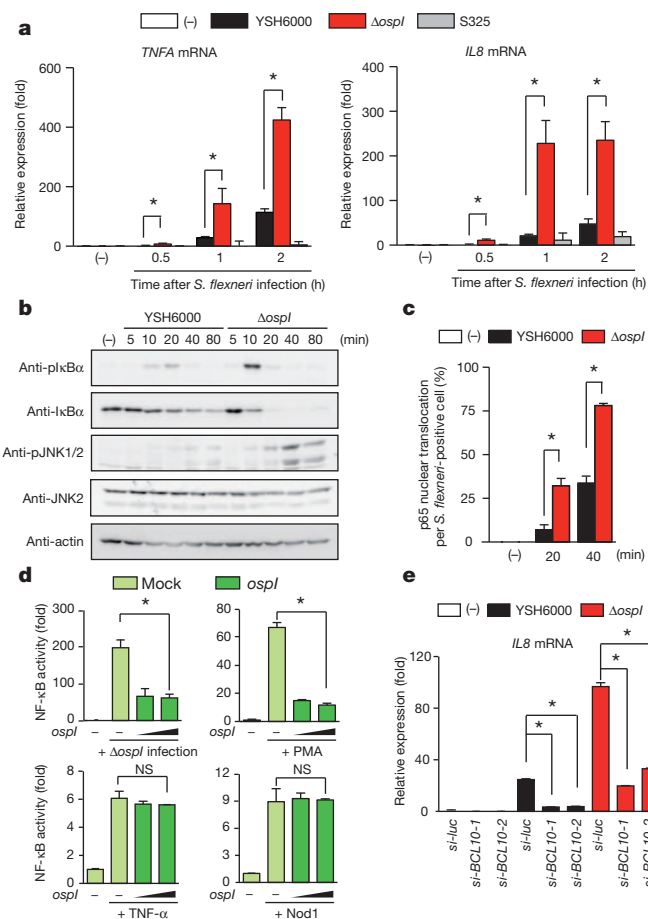


Figure 1 | OspI inhibits DAG-mediated NF- κ B activation during *S. flexneri* infection. **a**, The *TNFA* and *IL8* expression levels in HeLa cells at 0.5, 1 and 2 h post infection with YSH6000, Δ ospI or S325. The white bar (-) indicates no infection. **b**, Phosphorylation of I κ B α and JNK1/2 in HeLa cells infected with YSH6000 or Δ ospI at various times after infection, as detected with the indicated antibodies. pI κ B α , phosphorylated I κ B α ; pJNK1/2, phosphorylated JNK1/2. **c**, Nuclear localization of p65 in HeLa cells infected with YSH6000 or Δ ospI at 20 and 40 min post infection. **d**, The luciferase activity of NF- κ B at 90 min post infection is shown. NS, not significant. **e**, The *IL8* expression levels at 60 min post infection in HeLa cells transfected with BCL10-directed siRNA and infected with YSH6000 or Δ ospI. Data are presented as the mean \pm s.d.; $n = 3$; *, $P < 0.05$. si-luc, siRNA-luciferase.

interacts with BCL10 (Supplementary Fig. 4d). These results suggested that *S. flexneri*-induced changes in the host cell membrane stimulate the DAG-CBM-complex-NF- κ B pathway and that OspI specifically dampens this pathway.

To gain further structural and functional insight, we determined the crystal structure of recombinant *S. flexneri* OspI at 2.0 Å resolution (Supplementary Table 1; Protein Data Bank (PDB) ID 3B21). OspI has an α/β fold with four β -strands (β 1– β 4), seven α -helices (α 1– α 7) and one 3_{10} helix (Fig. 2a, b).

A search of known structures in PDB revealed that OspI shares structural homology with a cysteine protease family and is most closely related to AvrPphB (Supplementary Fig. 5). AvrPphB is a *Pseudomonas syringae* effector that is delivered by way of a T3SS and has been classified into a superfamily of enzymes containing cysteine proteases, acetyl transferases, deamidases, transglutaminases and ubiquitin carboxy-terminal hydrolase^{15,16}. Although there is considerable divergence in the overall folding across this superfamily, a core anti-parallel β -sheet and an α -helix that contains the active site cysteine, which packs against the β -strands, are conserved across this family (Supplementary Fig. 6). A potential catalytic triad (C62, H145 and D160) in OspI was identified based on a comparison with the active site of

AvrPphB (Fig. 2c and Supplementary Figs 5 and 6a). Although overall amino acid sequence similarity was low between OspI and these proteins, superimposing H145 and D160 of OspI onto H212 and D227 of AvrPphB or other members of this superfamily revealed remarkable similarity^{17–21} (Fig. 2c).

C62 of OspI, however, existed in three discrete conformations in the crystal structure, and the S γ position was located on the opposite side of the active site in AvrPphB. The fractional occupancy of each conformer was estimated to be 0.55 (conformation A), 0.35 (conformation B) and 0.1 (conformation C). The highest occupancy site of C62 appeared to form a disulphide bond with C65 at 2.05 Å (Fig. 2c, d and Supplementary Fig. 7). Alternative conformations of a deamidase active site cysteine are also seen in cytotoxic necrotizing factor 1 (Cnfl)²².

To characterize the putative catalytic triad, we substituted C62, H145 and D160 with serine (C62S) or alanine (C62A, H145A and D160A), and the mutant constructs were tested for their ability to suppress NF- κ B activation. Complementing the Δ ospI mutant with plasmids carrying the ospI-C62S, ospI-H145A or ospI-D160A genes failed to mitigate the increased I κ B α phosphorylation and *IL8* induction induced by Δ ospI infection (Fig. 3a, b and Supplementary Fig. 8c). OspI-C62S, OspI-H145A and OspI-D160A lost the ability to suppress NF- κ B activity on *S. flexneri* infection or PMA stimulation (Supplementary Fig. 8a, b), indicating that C62, H145 and D160 in OspI compose the catalytic triad that suppresses NF- κ B signalling.

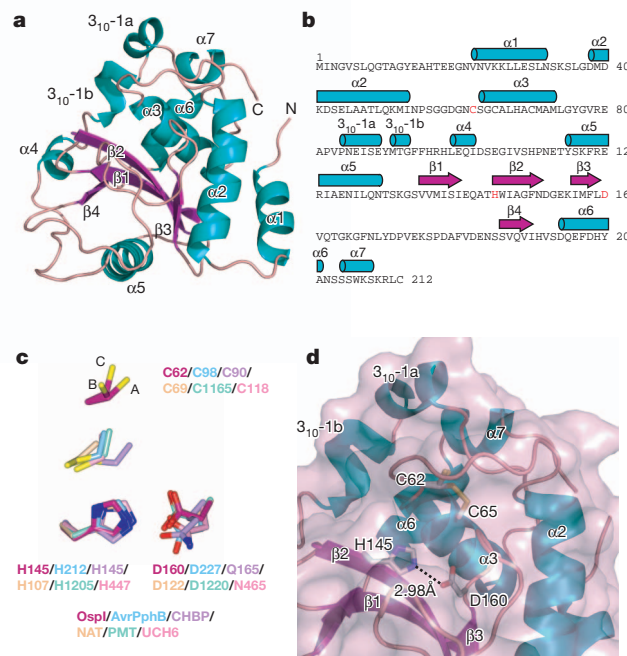


Figure 2 | Crystal structure of *S. flexneri* OspI. **a**, Overall structure of *S. flexneri* OspI. The secondary structure elements are coloured as follows: α -helices, cyan; β -strands, dark pink; and loops, light pink. **b**, The amino acid sequence of OspI. Secondary structure elements are coloured as in **a**. The putative active site residues are shown in red. **c**, Alignments of the catalytic cores of OspI with *P. syringae* AvrPphB and other members of the superfamily to which AvrPphB belongs, using the putative catalytic triad residues. All atoms of histidine and the main chain atoms of aspartic acid are shown as a reference. Shown are *S. flexneri* OspI, AvrPphB (PDB ID 1UKF), CHBP (Cif homologue from *Burkholderia pseudomallei*; PDB ID 3GQM), NAT (arylamine N-acetyltransferase from *Salmonella* Typhimurium; PDB ID 1E2T), PMT (*Pasteurella multocida* toxin; PDB ID 2EBF), and UCH6 (ubiquitin C-terminal hydrolase 6; PDB ID 1VJV). C62 in OspI is represented in three alternative conformations with the three conformers labelled A, B and C. **d**, A close-up view of the putative active site of OspI. OspI is shown with a surface representation. The putative active site residues are shown as stick models. Hydrogen bonding is indicated by a dashed line.

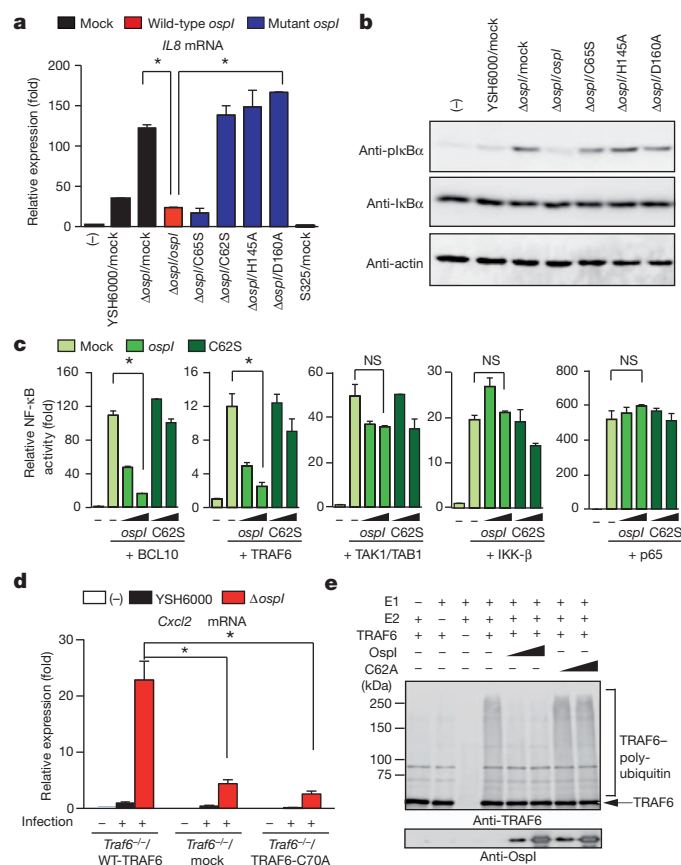


Figure 3 | Involvement of the CHD triad in inhibiting TRAF6-dependent NF- κ B activation. **a, b**, IL8 induction (at 60 min) and I κ B α phosphorylation (at 10 min) in HeLa cells infected with Δ ospI harbouring the indicated ospI mutant plasmids. **c**, NF- κ B luciferase reporter assay in HeLa cells transfected with the indicated plasmid combinations. **d**, Tra6-deficient MEFs were rescued with the indicated retroviruses. Cxcl2 induction in cells infected with YSH6000 or Δ ospI is shown. **e**, In vitro TRAF6 self-ubiquitylation assay in the presence of wild-type OspI or OspI-C62A (at 0.08 or 0.8 μ M, as indicated by the wedges). Data are presented as the mean \pm s.d.; $n = 3$; *, $P < 0.05$.

We determined which steps in the DAG-CBM-complex-NF- κ B pathway are targeted by OspI by examining NF- κ B activity in HeLa/OspI or HeLa/OspI-C62S cells that were transfected with a vector expressing the gene encoding BCL10, TRAF6, TAK1/TAB1, IKK- β or NF- κ B (p65). OspI but not OspI-C62S suppressed NF- κ B activity when HeLa cells were transfected with the gene encoding BCL10 or TRAF6 but not TAK1/TAB1, IKK- β or p65, suggesting that OspI targets TRAF6 or an upstream step in the signalling pathway (Fig. 3c). Thus, we tested rescued Tra6-deficient mouse embryonic fibroblasts (Tra6^{-/-} MEFs) for a possible role of OspI in modulating TRAF6 activation during *S. flexneri* infection: the MEFs tested were Tra6^{-/-}/WT-TRAF6 (Tra6^{-/-} MEFs expressing wild-type human TRAF6), Tra6^{-/-}/TRAF6-C70A (an E3-ligase-deficient mutant) and Tra6^{-/-}/mock. Cxcl2 levels in Δ ospI-infected Tra6-deficient MEFs were low and could be rescued by wild-type TRAF6, but not by TRAF6-C70A or mock treatment (Fig. 3d and Supplementary Fig. 8d). This result is consistent with OspI interference in TRAF6 activation during *S. flexneri* infection.

TRAF6 is an E3 ubiquitin ligase that cooperates with ubiquitin-activating E1 and ubiquitin-conjugating E2 enzymes (UBC13 and UEV1A), which are required for TRAF6 self-ubiquitylation and TRAF6-induced NF- κ B activation^{23–26}. Therefore, we investigated the effects of OspI and OspI-C62A on TRAF6 by using a self-ubiquitylation assay. We found that OspI, but not OspI-C62A, dampened TRAF6 polyubiquitylation (Fig. 3e and Supplementary Fig. 9). However, OspI did not affect the formation of E2-ubiquitin thioester intermediates

(Supplementary Fig. 10), suggesting that OspI modifies TRAF6, UBC13, UEV1A or ubiquitin. We incubated OspI with each of these putative targets and examined their electrophoretic mobility. We found that the mobility of UBC13, but not of the other proteins, shifted in an OspI-dose-dependent manner (Supplementary Fig. 11). Indeed, OspI could interact with His-UBC13 (Supplementary Fig. 12).

We then identified two affected peptides (WSPALQIR and DKWSPALQIR, which overlap each other) from a tryptic digest of UBC13 by using tandem mass spectrometry, and we found that the Q100 residue in UBC13 was deamidated to E100 by OspI (Fig. 4a and Supplementary Fig. 13). To confirm this, we created UBC13-Q100E and showed that this modification produced the same mobility shift as the modification at Q100 by OspI; treatment with OspI-C62A, OspI-H145A or OspI-D160A produced no mobility shift (Fig. 4b, c and Supplementary Fig. 14). Endogenous UBC13 was consistently deamidated in HeLa cells 10 min after infection with YSH6000 and Δ ospI/ospI but not Δ ospI or Δ ospI/ospI-C62S (Fig. 4d).

Importantly, the cocrystal structures of UBC13 and the zinc finger of TRAF6 indicated that Q100 of UBC13 was near the catalytic pocket but was also located near the interface between UBC13 and the TRAF6 zinc finger²⁷. Thus, we further characterized the ubiquitin-conjugating E2 activity of UBC13-Q100E using a ubiquitylation assay. The efficiency of UBC13-Q100E-catalysed ubiquitin chain formation on TRAF6 was markedly lower than that of wild-type UBC13 (Fig. 4e). In an NF- κ B reporter assay, UBC13-Q100E acted as a dominant negative because UBC13-Q100E suppressed the NF- κ B activity that was stimulated by PMA, TRAF6 and infection, but not TNF- α , in a dose-dependent manner (Supplementary Fig. 15). We thus concluded that OspI deamidates Q100 in UBC13, inhibiting the TRAF6-NF- κ B pathway.

Here we identified OspI as a new T3SS effector that specifically targets TRAF6-dependent acute inflammatory signalling during *S. flexneri* invasion of epithelial cells. OspI selectively deamidates UBC13, severely impairing the E2 ubiquitin ligase activity of UBC13, which is required for TRAF6 polyubiquitylation. Thus, *S. flexneri* can block acute NF- κ B-mediated inflammatory responses at the early stage of epithelial invasion (Supplementary Fig. 1), suggesting that OspI is a unique T3SS effector

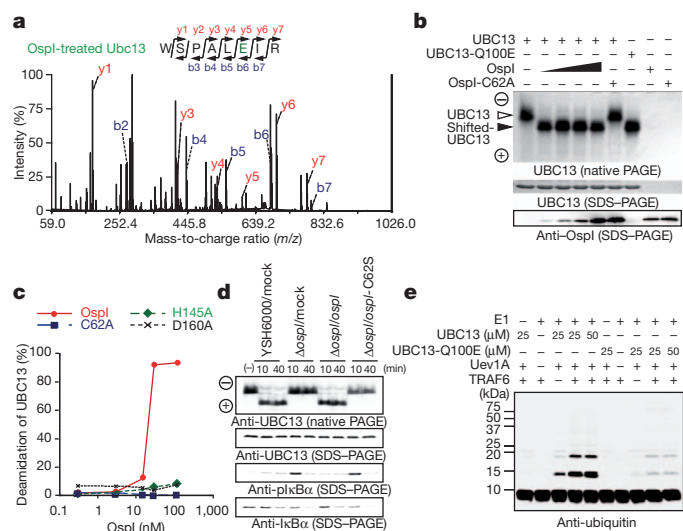


Figure 4 | OspI selectively deamidates Q100 in UBC13. **a**, Liquid chromatography and tandem mass spectrometry (LC-MS/MS) spectrum of a Q100-containing tryptic peptide in OspI-treated UBC13. **b**, Native PAGE and SDS-PAGE analysis of UBC13 treated with OspI (0.08, 0.8, 8 or 80 μ M) or OspI-C62A (80 μ M). **c**, Native PAGE analysis of UBC13 treated with OspI, OspI-C62A, OspI-H145A or OspI-D160A. **d**, HeLa cells were infected with *S. flexneri*/mock, Δ ospI/mock, Δ ospI/ospI or Δ ospI/ospI-C62S. At 10 and 40 min after infection, the cell lysates were subjected to native PAGE followed by anti-UBC13, anti-IκBα or anti-phospho-IκBα immunoblotting. **e**, In vitro ubiquitylation assay in the presence of UBC13 or UBC13-Q100E.

that dampens TRAF6-dependent inflammatory signalling in response to bacterial invasion of epithelial cells.

METHODS SUMMARY

The *S. flexneri* strains used in this study are described in the Methods. The procedures for the OspI crystallization and structure determination, mass spectrometric analysis, microarray analysis, quantitative PCR with reverse transcription, *in vitro* ubiquitylation assays, *in vitro* deamidation assays, luciferase assays, immunostaining, immunoblotting, *S. flexneri* infection and protein expression are described in the Methods, together with a description of the antibodies and the materials. OspI, reported as ORF169b (NCBI accession number CAC05849), is encoded by an *S. flexneri* virulence plasmid. The *S. flexneri* *ospI* gene was disrupted using a λ red recombinase-mediated recombination system as previously described²⁸. The Δ ospI strain was complemented by introducing plasmids (pWKS130) encoding *ospI*-C62S-myc, *ospI*-H145A-myc or *ospI*-D160A-myc. Bacterial infections were carried out as previously described²⁹. Briefly, HeLa cells were infected with the various *S. flexneri* strains at a multiplicity of infection (m.o.i.) of 100. For the afimbrial adhesin (Afa)-expressing *S. flexneri* strains, cells were infected at an m.o.i. of 10. Twenty minutes after infection, the plates were transferred into fresh medium containing 100 μ g ml⁻¹ gentamicin, to kill extracellular bacteria. Statistical significance was determined using Student's *t*-test.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 July 2011; accepted 24 January 2012.

Published online 11 March 2012.

- Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140**, 805–820 (2010).
- Kim, M. *et al.* Bacterial interactions with the host epithelium. *Cell Host Microbe* **8**, 20–35 (2010).
- Ashida, H., Ogawa, M., Kim, M., Mimuro, H. & Sasakawa, C. Bacteria and host interactions in the gut epithelial barrier. *Nature Chem. Biol.* **8**, 36–45 (2011).
- Taxman, D. J., Huang, M. T. & Ting, J. P. Inflammasome inhibition as a pathogenic stealth mechanism. *Cell Host Microbe* **8**, 7–11 (2010).
- Hilbi, H. Bacterial jailbreak sounds cellular alarm: phagosome membrane remnants trigger signaling. *Cell Host Microbe* **6**, 102–104 (2009).
- Dupont, N. *et al.* *Shigella* phagocytic vacuolar membrane remnants participate in the cellular response to pathogen invasion and are regulated by autophagy. *Cell Host Microbe* **6**, 137–149 (2009).
- Ashida, H. *et al.* Cell death and infection: a double-edged sword for host and pathogen survival. *J. Cell Biol.* **195**, 931–942 (2011).
- Mattoo, S., Lee, Y. M. & Dixon, J. E. Interactions of bacterial effector proteins with host proteins. *Curr. Opin. Immunol.* **19**, 392–401 (2007).
- Rahman, M. M. & McFadden, G. Modulation of NF- κ B signalling by microbial pathogens. *Nature Rev. Microbiol.* **9**, 291–306 (2011).
- Girardin, S. E. *et al.* CARD4/Nod1 mediates NF κ B and JNK activation by invasive *Shigella flexneri*. *EMBO Rep.* **2**, 736–742 (2001).
- Holden, N. S. *et al.* Phorbol ester-stimulated NF- κ B-dependent transcription: roles for isoforms of novel protein kinase C. *Cell. Signal.* **20**, 1338–1348 (2008).
- Shahnazari, S. *et al.* A diacylglycerol-dependent signaling pathway contributes to regulation of antibacterial autophagy. *Cell Host Microbe* **8**, 137–146 (2010).
- Oancea, E., Teruel, M. N., Quest, A. F. & Meyer, T. Green fluorescent protein (GFP)-tagged cysteine-rich domains from protein kinase C as fluorescent indicators for diacylglycerol signaling in living cells. *J. Cell Biol.* **140**, 485–498 (1998).
- Rawlings, D. J., Sommer, K. & Moreno-Garcia, M. E. The CARMA1 signalosome links the signalling machinery of adaptive and innate immunity in lymphocytes. *Nature Rev. Immunol.* **6**, 799–812 (2006).
- Shao, F., Merritt, P. M., Bao, Z., Innes, R. W. & Dixon, J. E. A *Yersinia* effector and a *Pseudomonas* avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. *Cell* **109**, 575–588 (2002).
- Zhu, M., Shao, F., Innes, R. W., Dixon, J. E. & Xu, Z. The crystal structure of *Pseudomonas* avirulence protein AvrPphB: a papain-like fold with a distinct substrate-binding site. *Proc. Natl Acad. Sci. USA* **101**, 302–307 (2004).
- Crow, A. *et al.* Crystal structures of Cif from bacterial pathogens *Photobacterium luminescens* and *Burkholderia pseudomallei*. *PLoS ONE* **4**, e5582 (2009).
- Sinclair, J. C., Sandy, J., Delgoda, R., Sim, E. & Noble, M. E. Structure of arylamine *N*-acetyltransferase reveals a catalytic triad. *Nature Struct. Biol.* **7**, 560–564 (2000).
- Kitadokoro, K. *et al.* Crystal structures reveal a thiol protease-like catalytic triad in the C-terminal region of *Pasteurella multocida* toxin. *Proc. Natl Acad. Sci. USA* **104**, 5139–5144 (2007).
- Orth, J. H. *et al.* *Pasteurella multocida* toxin activation of heterotrimeric G proteins by deamidation. *Proc. Natl Acad. Sci. USA* **106**, 7179–7184 (2009).
- Wang, Y. *et al.* MMDB: annotating protein sequences with Entrez's 3D-structure database. *Nucleic Acids Res.* **35**, D298–D300 (2007).
- Buetow, L., Flatau, G., Chiu, K., Boquet, P. & Ghosh, P. Structure of the Rho-activating domain of *Escherichia coli* cytotoxic necrotizing factor 1. *Nature Struct. Biol.* **8**, 584–588 (2001).
- Wang, C. *et al.* TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* **412**, 346–351 (2001).
- Orth, J. H. *et al.* Activation of the I κ B kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell* **103**, 351–361 (2000).
- Lamothe, B. *et al.* Site-specific Lys-63-linked tumor necrosis factor receptor-associated factor 6 auto-ubiquitination is a critical determinant of I κ B kinase activation. *J. Biol. Chem.* **282**, 4102–4112 (2007).
- Fukushima, T. *et al.* Ubiquitin-conjugating enzyme Ubc13 is a critical component of TNF receptor-associated factor (TRAF)-mediated inflammatory responses. *Proc. Natl Acad. Sci. USA* **104**, 6371–6376 (2007).
- Yin, Q. *et al.* E2 interaction and dimerization in the crystal structure of TRAF6. *Nature Struct. Mol. Biol.* **16**, 658–666 (2009).
- Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
- Ashida, H. *et al.* A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKK γ to dampen the host NF- κ B-mediated inflammatory response. *Nature Cell Biol.* **12**, 66–73 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Fukuda for matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) analysis. We thank the members of the Sasakawa laboratory for their advice. We are grateful to R. Whittier for critical reading of the manuscript. Diffraction data were collected at the Osaka University beamline BL44XU at SPring-8. This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas (23121525 (T.M.)), a Grant-in-Aid for Specially Promoted Research (23000012 (C.S.)), a Grant-in Aid for Young Scientists (A) (23689027 (M.K.)), several Grants-in-Aid for Young Scientists (B) (23790471 (M.O.), 23790472 (H.A.) and 22790403 (T.S.)), a Grant-in-Aid for Scientific Research (B) (23390102 (H.M.)), a Grant-in-Aid for Challenging Exploratory Research (23659220 (H.M.)), a Grant-in-Aid for Scientific Research on Priority Areas (18073003 (C.S.)) and the Japan Initiative for Global Research Network on Infectious Diseases (C.S.). Part of this work was supported by grants from the Naito Foundation (M.K. and H.M.), the Waksman Foundation of Japan (M.O.), the Yakult Bio-Science Foundation (M.O.), the Yakult Central Institute (C.S.) and The Hayashi Memorial Foundation for Female Natural Scientists (M.K.).

Author Contributions T.S. and T.M. designed and performed the experiments. M.K., H.M., M.S., H.A., A.O., T. Kobayashi and M.O. assisted with the experiments. J.G., Y.S. and J.I.I. gave advice regarding the design of the experiments and provided TRAF6 materials. T. Koyama and S.N. made antibodies. A.O. made the Δ ospI mutant. C.S. and T.M. wrote the paper.

Author Information Atomic coordinates and structure factors for the OspI structure have been deposited in the PDB under ID 3B21. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.S. (sasakawa@ims.u-tokyo.ac.jp) and T.M. (mizushi@sci.u-hyogo.ac.jp).

METHODS

Strains and plasmids. *S. flexneri* strain YSH6000 (ref. 30) was used as the wild-type strain. S325 (*mxiA::Tn5*)³¹ was used as a T3SS-deficient negative control. The vectors pCMV-FLAG-IκB kinase-β (IKK-β), pCMV-FLAG-transforming growth factor-β-activated kinase 1 (TAK1), and pCMV-FLAG-TAK1-binding protein 1 (TAB1) were from G. Takaesu³². TRAF6 (ref. 33) was ligated into the plasmids pGEX6P-1, pCMV-FLAG and pMx-puro. Human *MALTI* and human *BCL10* cDNAs were ligated into the pCMV-FLAG plasmid. Site-directed mutagenesis of *ospI*, *UBC13* and *TRAF6* was performed using a QuikChange Site-Directed Mutagenesis Kit (Stratagene).

Antibodies and materials. Antiserum specific for OspI was obtained by immunizing rabbits with the *S. flexneri* OspI peptide (189–207 amino acids, VIHVSQDEFDHYANSSSWK) coupled to keyhole limpet haemocyanin using *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester. Anti-*S. flexneri* lipopolysaccharide was prepared as previously described³⁴. Anti-phospho-IκBα (Cell Signaling Technology), anti-IκBα (BD Transduction Laboratories), anti-phospho-JNK1/2 (Cell Signaling Technology), anti-JNK2 (Santa Cruz Biotechnology), anti-p65 (F-6; Santa Cruz Biotechnology), anti-M2 Flag (Sigma), anti-His (Sigma), anti-actin (Millipore), anti-UBC13 (Invitrogen) and anti-ubiquitin (1B3; MBL) antibodies, as well as Alexa Fluor 633 phalloidin (Invitrogen), were obtained commercially and used as primary antibodies for immunostaining and immunoblotting. TNF-α (Peprotech), PMA (Sigma) and propranolol (Sigma) were obtained commercially.

Plasmid construction, expression and protein purification for crystallization. To purify OspI, the corresponding DNA was subcloned into the pGEX6P-1 glutathione S-transferase (GST) fusion vector (GE Healthcare). The plasmids were used to transform the BL21 strain of *Escherichia coli*, and protein expression was induced by adding 0.1 μM isopropyl-β-D-thiogalactoside (IPTG) at 30 °C for 6 h. Whole-cell extracts were prepared by sonicating and incubating the bacteria in lysis buffer (20 mM Tris-HCl, pH 7.4, and 150 mM NaCl) for 30 min with rotation at 4 °C, and then clarified by centrifugation at 27,000g for 20 min. GST fusion proteins were purified by incubating the clarified cell extract with Glutathione Sepharose 4B. The GST moiety was removed by cleaving with the PreScission Plus protease (GE Healthcare) and then performing anion-exchange chromatography (HiTrap Q FF, GE Healthcare) and gel-filtration chromatography (Superdex 75, GE Healthcare). A selenomethionine (Se-Met) derivative of OspI was prepared by expressing the protein in strain B834 *E. coli* cells that were grown in minimal medium supplemented with L-selenomethionine (25 μg ml⁻¹) (Calbiochem).

Crystallization. Native and Se-Met OspI were concentrated to 20 mg ml⁻¹ by ultrafiltration in 25 mM Tris-HCl, pH 7.5, and 1 mM dithiothreitol. Crystals were grown by the hanging-drop vapour diffusion method at 288 K in drops containing a mixture of 2 μl of protein solution and 2 μl of reservoir solution. For native and Se-Met OspI, the reservoir solution consisted of 0.1 M MES, pH 6.5, 0.1 M sodium acetate, 1.0 M NaCl and 24% (w/v) PEG 8000. The crystals of native and Se-Met OspI were transferred into a cryoprotective solution containing 0.1 M MES, pH 6.5, 0.1 M sodium acetate, 1.0 M NaCl and 35% (w/v) PEG 8000.

Data collection. X-ray diffraction data were collected at Spring-8 at 100 K in BL44XU. Single-wavelength anomalous diffraction (SAD) data were collected at a wavelength of 0.9789 Å. Native OspI crystal data were collected at a wavelength of 0.9000 Å. All data were reduced using the software Denzo, Scalepack³⁵ and programs from the CCP4 package³⁶. The data-processing statistics are provided in Supplementary Table 1.

Structure determination and refinement. The structure of OspI was determined using the SAD method. The positions of the heavy atoms were obtained using the software SHELXD³⁷ and refined using MLPHARE. The initial SAD phases were extended to a higher resolution with diffraction data that were collected from the native crystal to 2.0 Å resolution, with solvent flattening using PIRATE and DM³⁸. The initial model was constructed with the ARP/wARP³⁸ program. The remaining structural elements were built manually using the Coot³⁹ program. The model was refined to 2.0 Å resolution with the software REFMAC5 and CNS⁴⁰. The final OspI model shows amino acids 21 to 212. The phasing and refinement statistics are summarized in Supplementary Table 1. There are no residues in the disallowed regions of the Ramachandran plot. The structure figures were generated using CCP4mg⁴¹ and PyMOL⁴².

Cell culture and retroviral expression. HeLa and 293T cells were cultured in minimal essential medium (Sigma) and Dulbecco's modified Eagle's medium (Sigma), respectively, containing 10% FBS. *Traf6*-deficient MEFs were as described previously³⁴. Human TRAF6 and the C70A mutant were ligated into pMx-puro. The plasmids were transfected into Plat-E packaging cells as previously described³⁴. The resultant retrovirus was used to infect *Traf6*-deficient MEFs and selected for puromycin (2 μg ml⁻¹) resistance.

Luciferase reporter assay. For the luciferase reporter assays, cells were transfected with the reporter plasmid (pGL4.32[luc2P/NF-κB-RE/Hygro]; Promega), *Renilla* luciferase constructs (pRL-TK; Promega) and various combinations of expression

plasmids, by using FuGENE6 transfection reagent (Roche). The cells were infected with *S. flexneri* (m.o.i. of 20) or treated with PMA (50 nM) and TNF-α (2 ng ml⁻¹) for 1.5 h.

siRNA knockdown analysis. HeLa cells were transfected with siRNA (RNAi Co) by reverse transfection using Lipofectamine RNAiMAX. The following siRNAs were used: human *BCL10* siRNA-1 (5'-CCUUAAGAUCACGUACUGUUU-3' and 5'-ACAGUACGUGAUCUUAAGGG-3'); *BCL10* siRNA-2 (5'-CGUACUGUUUACGACAAUGA-3' and 5'-AUUGUCGUGAAACAGUACGUG-3'); and *si-luc*, 5'-CGUACGCGGAAUACUUCGAdTdT-3' and 5'-UCGAAGUAU UCCGCGUACGdTdT-3'.

RNA extraction and quantitative RT-PCR analysis. Total RNA was extracted using ISOGEN (Nippongene). First-strand cDNA was synthesized from 1 μg total RNA with reverse transcriptase using oligo(dT) primers. Real-time PCR was performed on cDNA samples using a LightCycler 2.0 (Roche) with the SYBR Green system (TaKaRa). The *GAPDH* expression levels were evaluated as an internal control. The following primer pairs were used: human *IL8* (5'-CTGATTTCTGCAGCTCTGTGTG-3' and 5'-GTCCACTCTCAATCACTCTCAG-3'), human *TNFA* (5'-CTTCAGACACCCTCAACCTCTT-3' and 5'-CACATTCTTGAATCCCAAGGT-3'), human *CCL20* (5'-TTGATGTCAGTGTGCTGCTACTCC-3' and 5'-CCGTGTGAAGCCACAATA-3'), human *CCL2* (5'-GCTCATAGCA GCCACCTTCATT-3' and 5'-CAGCTTCTTTGGGACACTTGCT-3'), human *CXCL2* (5'-GGGTGGCAAAGAAAAGGAG-3' and 5'-GTTGAGCGTCAAGAC CCAGT-3'), human *IL6* (5'-GATGGCTGAAAAGATGGATGC-3' and 5'-CTGCAGGAAGTGGATCAGGACT-3'), human *GAPDH* (5'-TGCCCTCAACGACCACTTTG-3' and 5'-TTCCTCTGTGTGCTCTTGCTGGG-3'), mouse *Cxcl2* (5'-CAAGGGTTGACTTCAAGAATCC-3' and 5'-CCTTGAGAGTGGCTA TGACTTC-3'), mouse *Gapdh* (5'-GTGTCTTACCACCATGGAG-3' and 5'-TCGTGGTTCACACCCATCAC-3') and human *BCL10* (5'-AGCGCGACC ATCGGAGAGGT-3' and 5'-TGTGGCCGAGAAATGGCAGG-3').

Purification of recombinant proteins and *in vitro* ubiquitylation assay. E1 (Boston Biochem) and ubiquitin (Sigma) were obtained commercially. His-UEV1A, His-UBC13 and GST-TRAF6 were prepared as previously described⁴⁴. E1 (200 ng), His-UEV1A (400 ng), His-UBC13 (500 ng) and TRAF6 (300 ng) were incubated at 30 °C for 20 min in a 50 μl reaction mixture containing reaction buffer (20 mM Tris-HCl, pH 7.5, 5 mM ATP, 5 mM MgCl₂ and 0.1 mM dithiothreitol (DTT)).

***In vitro* deamidation assay and mass spectrometric analysis.** UBC13 or ubiquitin (12 μg) was incubated with the indicated amount of purified OspI at 30 °C for 30 min in reaction buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl and 0.1 mM DTT). The reaction samples were separated by native PAGE, stained with Coomassie Brilliant Blue, and then quantified with the program ImageJ. The protein bands were excised from the gel, and each gel slice was subjected to in-gel digestion with trypsin. The resultant tryptic peptides were injected into a nano-liquid chromatography system (DiNa, KYA). The eluent was mixed with the matrix solution, and the mixture was directly blotted onto a matrix-assisted laser desorption (MALDI) sample plate with a micro-fractionation system (AccuSpot, Shimadzu Biotech) for MALDI-time of flight (MALDI-TOF) mass spectrometry (MS) analysis. Overall peptide identification was performed using a MALDI-TOF/TOF 4600 proteomics analyser (Applied Biosystems), followed by a database search with Mascot (Matrix Science).

Immunofluorescence analysis. HeLa cells were seeded on cover slips in six-well plates. The cells were infected with YSH6000 (wild-type *S. flexneri*) (m.o.i. of 100) for the indicated time periods, washed with PBS and then fixed in 4% paraformaldehyde in PBS for 15 min. The cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min and then blocked with 1% BSA in TBS for 30 min.

Congo-red-induced secretion. Congo-red-induced secretion of type III effectors was performed as reported previously⁴⁵. Briefly, *S. flexneri* incubated at 37 °C for 4 h was stimulated with 0.003% Congo red for 10 min at 37 °C. The supernatant was recovered by centrifugation. The samples were separated by SDS-PAGE and immunoblotted with anti-MYC antibody.

Microarray analysis. Total RNA was prepared 1 h after infection and analysed by a gene microarray (GeneChip Human Genome U133 Plus 2.0 Array, Affymetrix). The raw data were normalized and analysed using GeneSpring software. Fold changes in the gene expression values (relative to uninfected cells) were log₂-transformed and visualized as a heat map, indicated by the colour scale.

30. Sasakawa, C. *et al.* Molecular alteration of the 140-megadalton plasmid associated with loss of virulence and Congo red binding activity in *Shigella flexneri*. *Infect. Immun.* **51**, 470–475 (1986).
31. Sasakawa, C., Makino, S., Kamata, K. & Yoshikawa, M. Isolation, characterization, and mapping of Tn5 insertions into the 140-megadalton invasion plasmid defective in the mouse Sereny test in *Shigella flexneri* 2a. *Infect. Immun.* **54**, 32–36 (1986).

32. Takaesu, G. *et al.* TAB2, a novel adaptor protein, mediates activation of TAK1 MAPKKK by linking TAK1 to TRAF6 in the IL-1 signal transduction pathway. *Mol. Cell* **5**, 649–658 (2000).
33. Ishida, T. *et al.* Identification of TRAF6, a novel tumor necrosis factor receptor-associated factor protein that mediates signaling from an amino-terminal domain of the CD40 cytoplasmic region. *J. Biol. Chem.* **271**, 28745–28748 (1996).
34. Kim, M. *et al.* Bacteria hijack integrin-linked kinase to stabilize focal adhesions and block cell detachment. *Nature* **459**, 578–582 (2009).
35. Otwinowski, Z. M. W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
36. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
37. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
38. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* **374**, 229–244 (2003).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
40. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
41. McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. D* **67**, 386–394 (2011).
42. DeLano, W. L. Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol.* **12**, 14–20 (2002).
43. Kobayashi, N. *et al.* Segregation of TRAF6-mediated signaling pathways clarifies its role in osteoclastogenesis. *EMBO J.* **20**, 1271–1280 (2001).
44. Morikawa, H. *et al.* The bacterial effector Cif interferes with SCF ubiquitin ligase function by inhibiting deneddylation of Cullin1. *Biochem. Biophys. Res. Commun.* **401**, 268–274 (2010).
45. Parsot, C., Menard, R., Gounon, P. & Sansonetti, P. J. Enhanced secretion through the *Shigella flexneri* Mxi-Spa translocon leads to assembly of extracellular proteins into macromolecular structures. *Mol. Microbiol.* **16**, 291–300 (1995).

Tissue factor and PAR1 promote microbiota-induced intestinal vascular remodelling

Christoph Reinhardt^{1,2,3}, Mattias Bergentall^{1,2}, Thomas U. Greiner^{1,2}, Florence Schaffner⁴, Gunnel Östergren-Lundén^{1,2}, Lars C. Petersen⁵, Wolfram Ruf⁴ & Fredrik Bäckhed^{1,2,6}

The gut microbiota is a complex ecosystem that has coevolved with host physiology. Colonization of germ-free (GF) mice with a microbiota promotes increased vessel density in the small intestine¹, but little is known about the mechanisms involved. Tissue factor (TF) is the membrane receptor that initiates the extrinsic coagulation pathway², and it promotes developmental and tumour angiogenesis^{3,4}. Here we show that the gut microbiota promotes TF glycosylation associated with localization of TF on the cell surface, the activation of coagulation proteases, and phosphorylation of the TF cytoplasmic domain in the small intestine. Anti-TF treatment of colonized GF mice decreased microbiota-induced vascular remodelling and expression of the proangiogenic factor angiopoietin-1 (Ang-1) in the small intestine. Mice with a genetic deletion of the TF cytoplasmic domain or with hypomorphic TF (*F3*) alleles had a decreased intestinal vessel density. Coagulation proteases downstream of TF activate protease-activated receptor (PAR) signalling implicated in angiogenesis⁵. Vessel density and phosphorylation of the cytoplasmic domain of TF were decreased in small intestine from PAR1-deficient (*F2r*^{-/-}) but not PAR2-deficient (*F2rl1*^{-/-}) mice, and inhibition of thrombin showed that thrombin–PAR1 signalling was upstream of TF phosphorylation. Thus, the microbiota-induced extravascular TF–PAR1 signalling loop is a novel pathway that may be modulated to influence vascular remodelling in the small intestine.

The mammalian intestine is an organ with marked postnatal vascular adaptation, which is induced at weaning and coincides with the development of an adult microbiota. In agreement with early studies showing that the gut microbiota affects vascular remodelling in the intestine¹, we showed significant increases in villus width in the small intestine of conventionally raised (CONV-R) mice in comparison with GF mice (Fig. 1a), suggesting a link between vascular remodelling and altered villus architecture on colonization. We also showed increased staining and messenger RNA levels of the vascular marker platelet–endothelial cell adhesion molecule 1 (PECAM-1) in the small intestine of both CONV-R and conventionalized (CONV-D; GF mice that had been colonized for 14 days with a normal microbiota from a CONV-R mouse) mice in comparison with GF mice (Fig. 1b–d). The increased vessel density was located to the mid-distal part of the small intestine (Supplementary Fig. 1). Staining for the tip-cell marker delta-like ligand 4 (Dll4)⁶ indicated that colonization initially promoted sprouting angiogenesis but that the number of tip cells returned to basal levels once villus remodelling was complete (Fig. 1e).

We found increased levels of mRNA for Ang-1 as well as increased phosphorylation of the Ang-1 receptor Tie-2 in the small intestine of CONV-R in comparison with GF mice (Fig. 1f,g), thus providing a potential mechanism for microbiota-induced vascular remodelling. Consistent with increased vessel density, vascular endothelial growth factor receptor 1 (VEGFR-1) expression was also higher in CONV-R mice, but there were no changes in any other components of the VEGF

pathway (Supplementary Fig. 2). The Ang-1–Tie-2 axis promotes the remodelling and sprouting of blood vessels^{7,8}. To confirm a role for Ang-1 in microbiota-induced vascular remodelling, we injected GF mice with the specific Ang-1 inhibitor mL4-3 before and during a 14-day colonization with a normal gut microbiota and showed decreases in Tie-2 phosphorylation and intestinal vessel density (Fig. 1h–j). We identified the epithelium as a source of Ang-1, because its expression was increased in isolated primary enterocytes from CONV-R mice in comparison with those from GF mice (Fig. 1k).

Angiogenesis is linked to the cellular initiation of coagulation, and TF signalling has been shown to modulate angiogenesis^{3,4}. Because bacterial components are known to stimulate the coagulation system⁹, we speculated that TF could have a function in microbiota-induced angiogenesis in the intestine. In agreement with earlier studies of TF localization in humans¹⁰ and mice¹¹, we identified TF predominantly in enterocytes of the villi of small intestine in both GF and CONV-R mice (Supplementary Fig. 3). We injected GF mice with anti-TF antibody or control IgG before and at 4 and 9 days after colonization with a normal caecal microbiota. Tissues were harvested 14 days after colonization, and we confirmed that the injected antibodies localized to the small intestine (Supplementary Fig. 4). Anti-TF treatment did not affect PECAM-1 staining in GF mice that were not colonized (Fig. 1l,m); neither were levels of VEGF-A, VEGFR-2 or VEGFR-3 mRNA in CONV-D mice affected (Supplementary Fig. 5). However, anti-TF treatment decreased villus width (Fig. 1n), vessel density (Fig. 1o,p) and expression of PECAM-1 and Ang-1 mRNA (Fig. 1q,r) in CONV-D mice, suggesting that TF promotes microbiota-induced remodelling of the villus vasculature. Similarly, the vessel density was decreased in intestines from mice expressing low levels of human TF (low-TF mice)¹² compared with mice expressing normal levels of human TF from a knocked-in minigene¹³ (Supplementary Fig. 6). Because neither humanized mouse strain expressed alternatively spliced TF, intestinal vascular remodelling seems to be independent of alternatively spliced TF¹⁴.

Paneth cells have been suggested to regulate microbiota-induced intestinal angiogenesis in mice, but they also have a large effect on angiogenesis independently of colonization status¹. Anti-TF treatment did not decrease the number of Paneth cells or mRNA levels of Paneth cell-derived cryptdin 2 in CONV-D mice (Supplementary Fig. 7), indicating that treatment with antibody has no cytotoxic effect on Paneth cells. In addition, vessel density was similar in colonized CR2-*tox176* transgenic mice, which lack Paneth cells¹⁵, and their wild-type littermates after treatment with anti-TF (Supplementary Fig. 8).

Next we investigated whether intestinal TF expression and activity differed between GF and CONV-R mice. We did not observe any differences in intestinal levels of mRNA for TF from the two groups of mice (Fig. 2a). In contrast, immunoblot analyses identified two TF-reactive bands, one with an apparent molecular mass of 33 kDa that

¹Sahlgrenska Center for Cardiovascular and Metabolic Research/Wallenberg Laboratory, University of Gothenburg, 413 45 Gothenburg, Sweden. ²Department of Molecular and Clinical Medicine, University of Gothenburg, 413 45 Gothenburg, Sweden. ³Center for Thrombosis and Hemostasis (CTH), University Medical Center Mainz, Experimental Research, 55131 Mainz, Germany. ⁴Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, California 92037, USA. ⁵Haemostasis Biology, Novo Nordisk A/S, DK-2760 Maalov, Denmark. ⁶Novo Nordisk Foundation Center for Basic Metabolic Research, Section for Metabolic Receptology and Enteroendocrinology, Faculty of Health Sciences, University of Copenhagen, Copenhagen, DK-2200, Denmark.

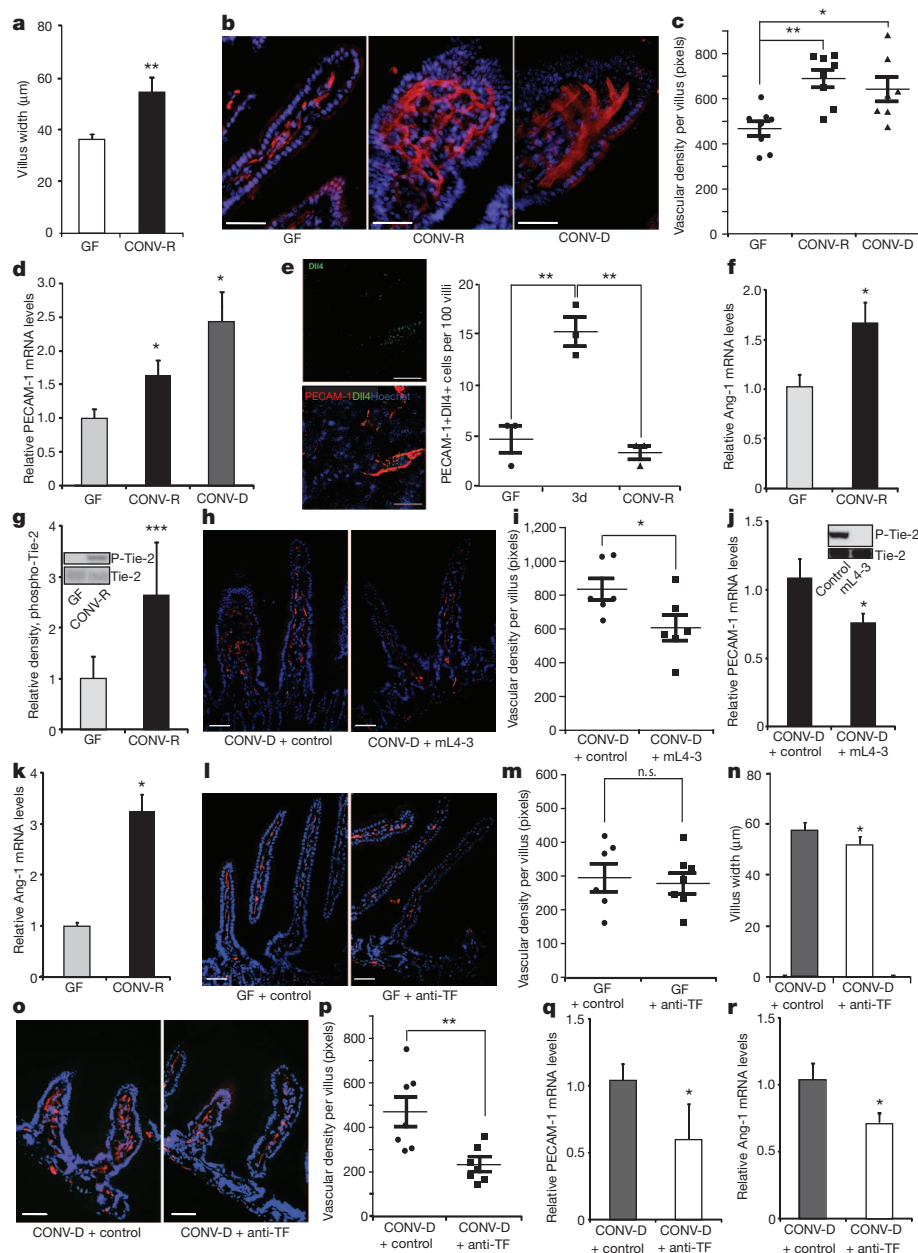


Figure 1 | TF promotes microbe-induced vascular remodelling in the gut.

a, Villus width of sections of small intestine from GF and CONV-R mice ($n = 4$ mice per group). **b**, PECAM-1 staining (red) of sections of small intestine from GF, CONV-R and CONV-D mice. Nuclei were stained with Hoechst nuclear dye (blue). **c**, Quantification of **b** ($n = 7$ or 8 mice per group). **d**, Relative levels of mRNA for the vascular marker PECAM-1 in GF, CONV-R and CONV-D mice ($n = 6$ or 7 mice per group). **e**, Dll4 staining (green) of sections of small intestine from GF, ex-GF mice colonized for 3 days (3d), and CONV-R mice. Endothelial cells were stained with PECAM-1 (red). Dll4-positive endothelial cells per 100 villi were quantified ($n = 3$ mice per group). **f**, Relative levels of mRNA for Ang-1 in sections of small intestine from GF and CONV-R mice ($n = 7$ –11 mice per group). **g**, Anti-phospho-Tie-2 immunoblot (Y1100 phosphorylation site) and quantification relative to total Tie-2 of small-intestinal lysates from GF and CONV-R mice ($n = 5$ mice per group). **h**, PECAM-1 staining of sections of small intestine from mice treated with control NaCl solution or the Ang-1 neutralizing peptide mL4-3. **i**, Quantification of **h** ($n = 6$ mice per group). **j**, Relative levels of mRNA for

PECAM-1 in sections of small intestine from CONV-D mice treated with NaCl control or mL4-3 ($n = 10$ or 11 mice per group). The inset shows that mL4-3 is a potent inhibitor of Ang-1-mediated Tie-2 phosphorylation. **k**, Relative levels of mRNA for Ang-1 in primary enterocytes from GF and CONV-R mice ($n = 10$ or 11 mice per group). **l**, PECAM-1 staining of sections of small intestine from GF mice treated with control or anti-TF antibody. **m**, Quantification of **l** ($n = 6$ or 7 mice per group). **n**, Villus width of sections of small intestine from CONV-D mice treated with control or anti-TF antibody ($n = 4$ mice per group). **o**, PECAM-1 staining of sections of small intestine from CONV-D mice treated with control or anti-TF antibody. **p**, Quantification of **o** ($n = 7$ mice per group). **q**, **r**, Relative levels of mRNA for PECAM-1 (**q**) and Ang-1 (**r**) in small intestine from CONV-D mice treated with control or anti-TF antibody ($n = 5$ or 6 mice per group). Female Swiss Webster mice or cells isolated from these mice were analysed in all panels. Scale bars, 50 μm. Results are shown as means \pm s.e.m. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.005$; n.s., not significant.

was present in both groups and a second band with an apparent molecular mass of 46 kDa that was present at higher levels in intestinal lysates from CONV-R and CONV-D mice (Fig. 2b, c). The gut microbiota has global effects on protein glycosylation in the small intestine¹⁶, which is necessary for the correct cellular localization and function of

many proteins including TF procoagulant activity¹⁷. We therefore speculated that the 46-kDa TF band resulted from microbiota-induced N-linked glycosylation of TF, the primary carbohydrate modification of TF¹⁷. In agreement with this, the mannose-binding lectin concanavalin A readily detected the 46-kDa form of TF in

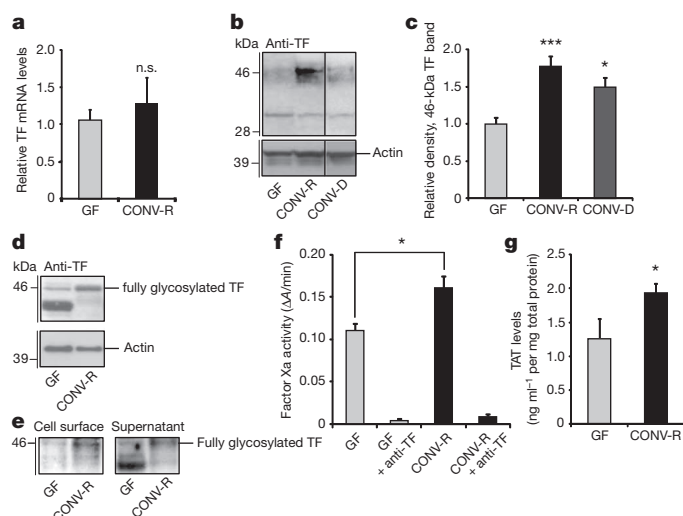


Figure 2 | The gut microbiota increases TF procoagulant activity and cell-surface localization. **a**, Relative levels of mRNA for TF in sections of small intestine from GF and CONV-R mice ($n = 7$ – 11 mice per group). **b**, Anti-TF immunoblot of small-intestinal lysates from GF, CONV-R and CONV-D mice. **c**, Quantification of the 46-kDa TF band shown in **b** ($n = 14$ – 25 mice per group). Data are normalized to actin and expressed relative to GF. **d**, Anti-TF immunoblot of primary enterocytes (from GF and CONV-R mice) after 2 h of culture. **e**, Anti-TF immunoblots from N-hydroxysuccinimido-biotin-labelled primary enterocytes from GF and CONV-R mice. Left: pull-down of proteins located on the plasma membrane with NeutrAvidin beads. Right: supernatant containing unlabelled proteins. **f**, Factor Xa activity in small-intestinal lysates from GF and CONV-R mice treated with control or anti-TF antibody (1H1; $n = 4$ or 5 mice per group). **g**, Levels of thrombin-antithrombin (TAT) complexes in small-intestinal lysates from GF and CONV-R mice ($n = 7$ mice per group). Female Swiss Webster mice or cells isolated from these mice were analysed in all panels. Results are shown as means \pm s.e.m. Asterisk, $P < 0.05$; three asterisks, $P < 0.005$; n.s., not significant.

small-intestinal lysates from CONV-D mice (Supplementary Fig. 9a). Treatment with the N-glycosidase PNGase F abolished detection of the 46-kDa form and generated a partly deglycosylated form with increased electrophoretic mobility that was only weakly detected by concanavalin A. We also treated primary enterocytes from CONV-R mice with the N-glycosylation inhibitor tunicamycin and observed a decreased abundance of the 46-kDa form (Supplementary Fig. 9b,c). These findings indicate that the gut microbiota promotes N-glycosylation of TF.

Exposure of functional TF on cell surfaces is regulated by basolateral sorting in epithelial cells¹⁸. Surface biotinylation followed by biotin pull-down of primary enterocytes from GF mice showed that the underglycosylated TF was mainly intracellular (Fig. 2d,e). In contrast, surface labelling of proteins or carbohydrates showed that enterocytes from CONV-R mice had high levels of the fully glycosylated TF on the cell surface (Fig. 2d,e and Supplementary Fig. 10). Confocal microscopy confirmed plasma membrane localization of TF in primary enterocytes isolated from CONV-R mice (Supplementary Fig. 11 and Supplementary Movie). These changes were associated with enhanced coagulation activation, as demonstrated by increased TF–FVIIa-dependent generation of coagulation factor Xa and higher levels of thrombin-antithrombin complexes in lysates of small intestine from CONV-R in comparison with those from GF mice (Fig. 2f,g).

Not only does TF initiate coagulation, it also interacts with integrins on the extracellular side and regulates integrin function through its cytoplasmic domain¹⁹. Proximity ligation and immunoprecipitation experiments showed increased TF– β_1 integrin complex formation in intestinal tissue from CONV-R mice in comparison with GF counterparts (Supplementary Fig. 12a,b). Furthermore, TF– β_1 integrin complex formation was decreased by treating CONV-R mice with

tunicamycin (Supplementary Fig. 12c). The cytoplasmic domain of TF contains a conserved Ser/Thr-Pro phosphorylation site²⁰. Phosphorylation of this domain has been observed at sites of neovascularization⁴; it requires surface localization of TF²¹ and regulates integrin function¹⁹. An antibody directed against the phosphorylated domain of mouse TF detected increased phosphorylation in the 46-kDa form of TF in lysates of small intestine from CONV-R in comparison with that from GF mice (Fig. 3a), and treatment of primary enterocytes from CONV-R mice with tunicamycin decreased levels of the phosphorylated 46-kDa form of TF (Fig. 3b,c).

To test directly whether the TF cytoplasmic domain was involved in vascular remodelling, we analysed small-intestinal tissue from mice with a targeted deletion of this domain (Δ CT mice)²² and age-matched wild-type mice. Δ CT mice had significantly decreased villus vascularization (Fig. 3d,e) and decreased expression of mRNA for PECAM-1 and Ang-1 in comparison with wild-type mice (Fig. 3f). TF from wild-type and Δ CT mice had similar electrophoretic mobilities (Supplementary Fig. 13). These data show that the TF cytoplasmic domain has a function in increasing vessel density in the small intestine but that it is not required for glycosylation. Anti-TF treatment decreased TF phosphorylation but not total TF levels in CONV-D mice (Supplementary Fig. 14). These results indicate that the inhibitory effects of anti-TF on vascular remodelling are independent of TF downregulation but, at least in part, involve inhibition of TF cytoplasmic domain phosphorylation.

TF also mediates signalling through coagulation proteases that activate the G-protein-coupled receptors PAR1 and PAR2 (refs 4, 5). We investigated the effect of the gut microbiota on PAR expression in small-intestinal tissue. Levels of mRNA for PAR1 but not those for PAR2 were increased in CONV-R mice in comparison with GF

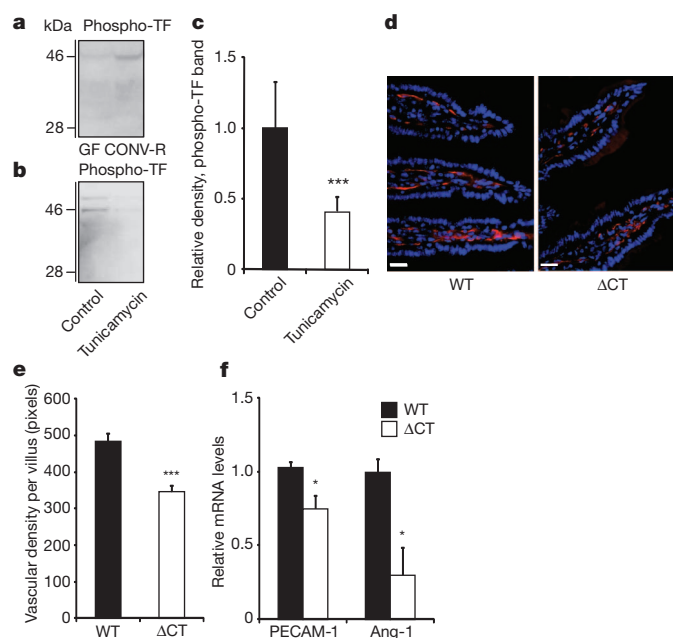


Figure 3 | The gut microbiota increases phosphorylation of the cytoplasmic tail of TF, which increases vessel density in the intestine. **a**, **b**, Anti-phospho-TF immunoblot of small-intestinal lysates from GF and CONV-R mice (**a**) and primary enterocytes (from CONV-R mice) incubated for 2 h in the absence and presence of tunicamycin ($10 \mu\text{mol l}^{-1}$) (**b**). **c**, Quantification of **b** ($n = 5$ mice per group). **d**, PECAM-1 staining (red) of sections of small intestine from 10–12-week-old wild-type (WT) and Δ CT female mice on a C57Bl6/J genetic background. Nuclei were stained with Hoechst nuclear dye (blue). **e**, Quantification of **d** ($n = 4$ – 6 mice per group). **f**, Relative levels of mRNA for PECAM-1 and Ang-1 in segments of small intestine from WT and Δ CT mice ($n = 3$ or 4 mice per group). Female Swiss Webster mice or cells isolated from these mice were analysed in **a**–**c**. Scale bars, $20 \mu\text{m}$. Results are shown as means \pm s.e.m. Asterisk, $P < 0.05$; three asterisks, $P < 0.005$.

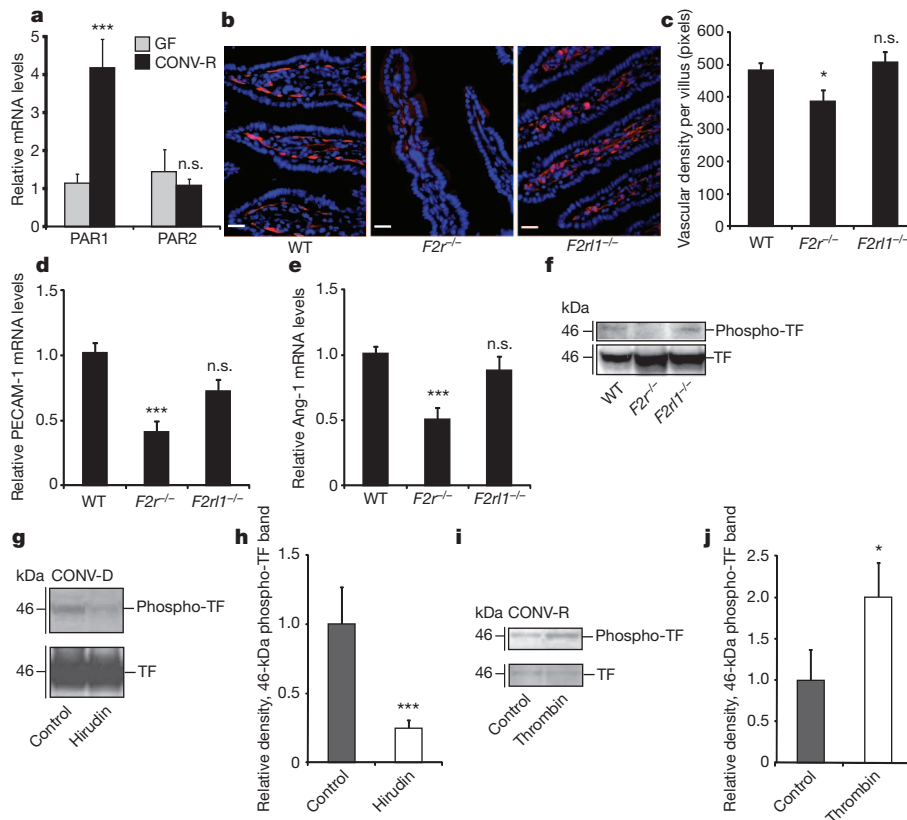


Figure 4 | PAR1 activation increases vessel density in the small intestine.

a, Relative levels of mRNA for PAR1 and PAR2 in segments of small intestine from GF and CONV-R mice ($n = 7$ or 8 mice per group). **b**, PECAM-1 staining (red) of sections of small intestine from wild-type (WT), *F2r*^{-/-} and *F2rl1*^{-/-} mice. Nuclei were stained with Hoechst nuclear dye (blue). **c**, Quantification of **b** ($n = 6$ – 9 mice per group). **d**, **e**, Relative levels of mRNA for PECAM-1 (**d**) and Ang-1 (**e**) in segments of small intestine from wild-type, *F2r*^{-/-} and *F2rl1*^{-/-} mice ($n = 6$ – 9 mice per group). **f**, **g**, Anti-TF and anti-phospho-TF immunoblots of small-intestinal lysates from WT, *F2r*^{-/-} and *F2rl1*^{-/-} mice (**f**) and CONV-D mice treated with PBS (control) or hirudin (1 mg/mouse)

counterparts (Fig. 4a). PAR1 is abundantly expressed in endothelial cells⁵, but we found that PAR1 was also expressed in enterocytes and at higher levels in cells from CONV-R mice than in GF counterparts (Supplementary Fig. 15). PECAM-1 staining (Fig. 4b,c) as well as levels of mRNA for PECAM-1 and Ang-1 (Fig. 4d,e) were decreased in intestinal tissue from PAR1-deficient (*F2r*^{-/-}) but not PAR2-deficient (*F2rl1*^{-/-}) mice, which is in agreement with a study showing that thrombin induces PAR1-dependent Ang-1 expression in endothelial cells²³. Together, these data show that the microbiota induces increased expression of PAR1, and that PAR1 has a role in remodelling the vasculature in the small intestine.

We next investigated the potential interrelation between PAR1 and TF in intestinal tissue. Phosphorylation of TF was decreased in lysates of small intestine from *F2r*^{-/-} in comparison with that from wild-type and *F2rl1*^{-/-} mice (Fig. 4f), indicating that PAR1 acts upstream of TF phosphorylation. We blocked thrombin and thrombin-dependent PAR1 signalling with hirudin immediately before and during colonization of GF mice for 6 h, and observed a striking decrease in TF phosphorylation in lysates of small intestine (Fig. 4g,h). We also showed that thrombin increased the phosphorylation of TF in primary enterocytes (Fig. 4i,j). Taken together, these data suggest that functional, procoagulant TF is required for the generation of thrombin, which in turn activates PAR1 to promote phosphorylation of the cytoplasmic domain of TF in enterocytes.

This study has uncovered a novel connection between TF, PAR1 and Ang-1 in modulating vascular remodelling after colonization. Our

immediately before colonization and at 2 h and 4 h after colonization

(**g**). **h**, Quantification of the phospho-TF band shown in **g** ($n = 6$ or 7 mice per group). **i**, Anti-TF and anti-phospho-TF immunoblots of primary enterocytes (from CONV-R mice) incubated for 2 h with human thrombin (50 nmol l⁻¹). **j**, Quantification of the phospho-TF band shown in **i** ($n = 8$ mice per group). Female Swiss Webster mice were analysed in **a** and **g**–**j**. Female WT, *F2r*^{-/-} and *F2rl1*^{-/-} mice on a C57BL/6/J genetic background were used in **b**–**f**. Scale bars, 20 μm. Results are shown as means ± s.e.m. Asterisk, $P < 0.05$; three asterisks, $P < 0.005$; n.s., not significant.

results support a model in which the microbiota induces increased glycosylation and surface localization of TF in the small intestine, leading to activation of coagulation, PAR1-dependent-phosphorylation of the TF cytoplasmic domain, and TF cytoplasmic domain signalling linked to Ang-1-dependent vascular remodelling (Supplementary Fig. 16). This pathway is distinct from established models of ocular angiogenesis⁴ or tumour-induced neovascularization, which requires the TF–Factor VIIa–PAR2-mediated induction of pro-angiogenic chemokines²⁴. We therefore suggest that TF may support distinct pro-angiogenic pathways in different tissues. Increased vascularization of the villi of the small intestine increases oxygenation of the villi, which are shortened and widened after colonization. This process may promote increased nutrient absorption, which has been associated with increased adiposity in CONV-R mice²⁵. Further dissection of how TF and PAR1 mediate postnatal microbiota-induced angiogenesis may provide new therapeutic targets for improving intestinal homeostasis and modulating the absorptive capacity of the gut.

METHODS SUMMARY

Mice. GF Swiss Webster female mice were maintained in flexible film isolators under a 12-h light cycle and fed with an autoclaved chow diet (Labdiet, St Louis) *ad libitum*. CONV-R Swiss Webster female mice were transferred into identical isolators at weaning. Age-matched female ΔCT (ref. 22), *F2r*^{-/-} (ref. 26) and *F2rl1*^{-/-} (ref. 27) mice and wild-type controls on a C57BL/6J background, low-TF and TF knock-in mice on a C57BL/6 background¹³, and transgenic CR2-*tox176* and non-transgenic littermates on a FVB/N genetic background¹⁵ were also used.

Mice were killed at 10–14 weeks of age by cervical dislocation or overdose anaesthesia; small intestines were removed and divided into eight equal segments. The fifth segment was used unless otherwise stated. Animal protocols were approved by the Research Animal Ethics Committee in Gothenburg and the Scripps Research Institute Institutional Animal Care and Use Committee (IACUC).

Isolation of primary enterocytes. Enterocytes were isolated from the fifth segment of small intestine as described previously²⁸. The cells were cultured for 2 h in DMEM medium before the beginning of the experiment.

Administration of TF antibody. Rabbit anti-mouse TF antibody²⁹ or rabbit anti-mouse IgG (Sigma) (1.33 mg per kg body weight) were administered intraperitoneally to GF mice before conventionalization with a caecal microbiota from a CONV-R donor. Additional antibody injections were given 4 and 9 days after colonization. The mice were killed 14 days after colonization.

Statistics. Data were analysed with Student's *t*-test for two sample groups or one-way analysis of variance for three sample groups.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 December 2009; accepted 23 January 2012.

Published online 11 March 2012.

- Stappenbeck, T. S., Hooper, L. V. & Gordon, J. I. Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc. Natl Acad. Sci. USA* **99**, 15451–15455 (2002).
- Morrissey, J. H., Fakhrai, H. & Edgington, T. S. Molecular cloning of the cDNA for tissue factor, the cellular receptor for the initiation of the coagulation protease cascade. *Cell* **50**, 129–135 (1987).
- Carmeliet, P. *et al.* Role of tissue factor in embryonic blood vessel development. *Nature* **383**, 73–75 (1996).
- Belting, M. *et al.* Regulation of angiogenesis by tissue factor cytoplasmic domain signaling. *Nature Med.* **10**, 502–509 (2004).
- Griffin, T. C., Srinivasan, Y., Zheng, Y.-W., Huang, W. & Coughlin, S. R. A role for thrombin receptor signaling in endothelial cells during embryonic development. *Science* **293**, 1666–1670 (2001).
- Hellström, M. *et al.* Dll4 signalling through Notch 1 regulates formation of tip cells during angiogenesis. *Nature* **445**, 776–780 (2007).
- Sato, T. N. *et al.* Distinct roles of the receptor tyrosine kinases Tie-1 and Tie-2 in blood vessel formation. *Nature* **376**, 70–74 (1995).
- Suri, C. *et al.* Requisite role of angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis. *Cell* **87**, 1171–1180 (1996).
- Iwanaga, S. The limulus clotting reaction. *Curr. Opin. Immunol.* **5**, 74–82 (1993).
- More, L. *et al.* Immunohistochemical study of tissue factor expression in normal intestine and idiopathic inflammatory bowel disease. *J. Clin. Pathol.* **46**, 703–708 (1993).
- Luther, T. *et al.* Tissue factor expression during human and mouse development. *Am. J. Pathol.* **149**, 101–113 (1996).
- Parry, G. C., Erlich, J. H., Carmeliet, P., Luther, T. & Mackman, N. Low levels of tissue factor are compatible with development and hemostasis in mice. *J. Clin. Invest.* **101**, 560–569 (1998).
- Snyder, L. A. *et al.* Expression of human tissue factor under the control of the mouse tissue factor promoter mediates normal hemostasis in knock-in mice. *J. Thromb. Haemost.* **6**, 306–314 (2008).
- van den Berg, Y. W. *et al.* Alternatively spliced tissue factor induces angiogenesis through integrin ligation. *Proc. Natl Acad. Sci. USA* **106**, 19497–19502 (2009).
- Garabedian, E. M., Roberts, L. J., McNevin, M. S. & Gordon, J. I. Examining the role of Paneth cells in the small intestine by lineage ablation in transgenic mice. *J. Biol. Chem.* **272**, 23729–23740 (1997).
- Bry, L., Falk, P. G., Midtvedt, T. & Gordon, J. I. A model of host–microbial interactions in an open mammalian ecosystem. *Science* **273**, 1380–1383 (1996).
- Krudysz-Amblo, J., Jennings, M. E. II, Mann, K. G. & Butenas, S. Carbohydrates and activity of natural and recombinant tissue factor. *J. Biol. Chem.* **285**, 3371–3382 (2010).
- Camerer, E. *et al.* Opposite sorting of tissue factor in human umbilical vein endothelial cells and Madin–Darby canine kidney epithelial cells. *Blood* **88**, 1339–1349 (1996).
- Dorflautner, A., Hintermann, E., Tarui, T., Takada, Y. & Ruf, W. Cross-talk of integrin $\alpha 3 \beta 1$ and tissue factor in cell migration. *Mol. Biol. Cell* **15**, 4416–4425 (2004).
- Zioncheck, T. F., Soumitra, R. & Vehar, G. A. The cytoplasmic domain of tissue factor is phosphorylated by a protein kinase C-dependent mechanism. *J. Biol. Chem.* **267**, 3561–3564 (1992).
- Dorflautner, A. & Ruf, W. Regulation of tissue factor cytoplasmic domain phosphorylation by palmitoylation. *Blood* **102**, 3998–4005 (2003).
- Melis, E. *et al.* Targeted deletion of the cytoplasmic domain of tissue factor in mice does not affect development. *Biochem. Biophys. Res. Commun.* **286**, 580–586 (2001).
- Blackburn, J. S. & Brinckerhoff, C. E. Matrix metalloproteinase-1 and thrombin differentially activate gene expression in endothelial cells via PAR-1 and promote angiogenesis. *Am. J. Pathol.* **173**, 1736–1746 (2008).
- Schaffner, F. *et al.* Cooperation of tissue factor cytoplasmic domain and PAR2 signaling in breast cancer development. *Blood* **116**, 6106–6113 (2010).
- Bäckhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl Acad. Sci. USA* **101**, 15718–15723 (2004).
- Connolly, A. J., Ishihara, H., Kahn, M. L., Farese, R. V. Jr & Coughlin, S. R. Role of the thrombin receptor in development and evidence for a second receptor. *Nature* **381**, 516–519 (1996).
- Damiano, B. P. *et al.* Cardiovascular responses mediated by protease-activated receptor-2 (PAR-2) and thrombin receptor (PAR-1) are distinguished in mice deficient in PAR-2 or PAR-1. *J. Pharmacol. Exp. Ther.* **288**, 671–678 (1999).
- Perreault, N. & Beaulieu, J.-F. Primary cultures of fully differentiated and pure human intestinal epithelial cells. *Exp. Cell Res.* **245**, 34–42 (1998).
- Furlan-Freguia, C., Marchese, P., Gruber, A., Ruggeri, Z. M. & Ruf, W. P2X7 receptor signaling contributes to tissue factor-dependent thrombosis in mice. *J. Clin. Invest.* **121**, 2932–2944 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Perkins for editing the manuscript; C. Arvidsson, A. Hallén, S. Wagoner, M. Karlsson, D. O'Donnell, S. Islam, N. Hörmann and A. Mohammadzadeh for technical assistance; A. Hallén for providing Supplementary Fig. 16; and P. Lindahl, J. Gordon, C. Betsholtz, M. Bergö, A. Wichmann, V. Tremaroli, M. Levin and S. Massberg for comments and suggestions. We are grateful to D. Kirchhofer for the gift of 1H1 monoclonal anti-mouse TF antibody, J. Nichols at Amgen for mL4-3, N. Mackman for the low-TF mice, M. Anderson for the human TF knock-in mice, and J. Gordon for providing CR2-*tox176* mice. This study was supported by the Swedish Foundation for Strategic Research, the Swedish Research Council, Torsten and Ragnar Söderberg's foundation, Petrus and Augusta Hedlund's foundation, and the Swedish federal government under the LUA/ALF agreement to F.B., National Institutes of Health grants HL-60742 and HL-77753 to W.R., and a Marie Curie Fellowship, a Marie Curie Reintegration Grant from the European Union and the German Federal Ministry of Education and Research (BMBF 01EO1003) to C.R.

Author Contributions C.R. was responsible for conception and study design, biochemical analysis of TF, analysis of vessel densities, data assembly and analysis, and writing the manuscript. M.B., T.U.G., F.S. and G.Ö.L. performed data collection, analysis and interpretation and commented on the manuscript. L.C.P. provided material. W.R. and F.B. were responsible for conception and study design, data analysis and interpretation, and writing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to F.B. (fredrik.backhed@wlab.gu.se).

METHODS

Administration of mL4-3. mL4-33 (ref. 30) (2.32 mg per kg body weight) was administered subcutaneously to GF mice before conventionalization with a caecal microbiota from a CONV-R donor. Additional injections of mL4-3 were given three times a week. The mice were killed 14 days after colonization.

Preparation of intestinal samples. For immunohistochemistry and *in situ* hybridization, the small intestine (divided into eight equal segments) and colon were flushed with PBS after excision and opened longitudinally. The tissue was fixed overnight in 4% formaldehyde at 4 °C, washed three times in PBS, and incubated in 10% sucrose in PBS at 4 °C. After 3 h, the buffer was replaced with 20% sucrose and 10% glycerol in PBS, and the tissue was incubated at 4 °C overnight. Tissues were dried with a paper towel and mounted in OCT on solid CO₂. Frozen sections 6 µm thick were prepared.

For mRNA analyses, the segments were frozen immediately at –80 °C in liquid nitrogen. For immunoblots, the fifth segment was flash-frozen and homogenized for 10 min in lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM EDTA, 1% Triton X-100) containing Roche Complete protease and PhosStop phosphatase inhibitors (diluted 1:10). The homogenate was incubated for 30 min on ice and centrifuged three times at 9,000g for 10 min to remove insoluble cell debris.

Immunohistochemistry. Sections were incubated for 20 min at 22 °C and blocked for 1 h with diluted TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton X-100) containing 5% rabbit serum. The blocking solution was removed and the following primary antibodies, diluted in the same blocking solution, were added: rat anti-mouse PECAM-1 (dilution 1:300; BD, Franklin Lakes), chicken anti-cytokeratin 8 (dilution 1:100; Abcam, Cambridge), goat anti-DLL4 (dilution 1:50; R&D), rabbit anti-mouse TF²⁹ (1 µg ml^{–1}) and rabbit anti-PAR1 (dilution 1:300; Sigma). The samples were incubated overnight at 4 °C, washed three times for 5 min in TBST and incubated for 1 h with secondary antibodies (Invitrogen, Carlsbad) at room temperature (rabbit anti-rat Alexa594, dilution 1:800; goat anti-rabbit IgG Alexa488, dilution 1:5,000; goat anti-chicken IgG Alexa488, dilution 1:2,000; all from BD). Nuclei were stained with Hoechst dye (3 µg ml^{–1}; Sigma) and the sections were washed three times for 10 min in TBST. For detection of Paneth cells, fluorescein isothiocyanate-isolectin (10 µg ml^{–1}; Sigma) was used. Slides were mounted, and viewed at ×20 and ×40 magnification with a fluorescence microscope (Axioplan 2 imaging; Zeiss, Oberkochen). Biopix IQ software (<http://www.biopix.se>) was used to quantify PECAM-1 staining in 2–11 villi per mouse. Confocal images and three-dimensional reconstructions were obtained with a Leica TCS SP5 confocal microscope (Leica, Wetzlar).

Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) analysis. Total RNA was isolated from small-intestinal tissues and isolated primary enterocytes with the RNeasy kit (Qiagen, Hilden). Total RNA (0.5 µg) was reverse transcribed (High Capacity cDNA Reverse Transcription kit; Applied Biosystems, Foster City) and SYBR green-based qRT-PCR was performed as described previously³¹. Primers are listed in Supplementary Table 1.

In situ hybridization. Mouse TF cDNA³² was subcloned into pSPT19 for subsequent *in vitro* RNA synthesis. Non-radioactive, digoxigenin-labelled sense and antisense RNA probes were synthesized with the DIG RNA Labelling Kit (SP6/T7; Roche, Mannheim). Tissues were pretreated for 2 min with proteinase K (10 µg ml^{–1}) in 50 mM Tris-HCl pH 7.5, 5 mM EDTA; the reaction was stopped by washing for 30 s in 0.2% glycine in PBS, followed by two additional washing steps in PBS. Tissues were fixed for 15 min in 4% paraformaldehyde in PBS, and washed in PBS for 2 min. Hybridization solution was added, and tissues were pre-hybridized for 1 h at 65 °C. RNA probe (8 ng µl^{–1} hybridization solution) was added and preheated for 5 min at 80 °C; 100 µl was added to each slide and incubated overnight at 65 °C in a humidified box. Slides were washed three times for 30 min in a preheated washing solution at 65 °C and twice for 30 min in MABT (100 mM maleic acid pH 7.5, 150 mM NaCl, 0.1% Tween 20) at room temperature. Slides were blocked with 2% blocking reagent (Roche, Mannheim), 20% heat-inactivated sheep serum (Sigma) in MABT for 1 h at room temperature. Binding of the RNA template was detected with alkaline-phosphatase-conjugated Fab fragments (Roche, Mannheim) and BM Purple.

Factor Xa activity. Factor Xa activity was measured in small-intestinal lysates as described previously³².

Measurement of TAT complexes. The TAT ELISA Kit (Uscnlife, Guangguguojia) was used for determination of the concentration of mouse TAT complexes in lysates of small-intestinal tissue.

Immunoprecipitation. Tissue lysates were incubated for 1 h with anti-mouse TF antibody (70 µg ml^{–1}; American Diagnostica, Stamford) or anti-integrin β₁ antibody (dilution 1:100 Cell Signaling, Danvers), and immunocomplexes were precipitated by adding 50 µl of Protein A-Sepharose fast flow 4B (Sigma). TF and integrin β₁ antigen were detected as described below.

Glycosidase treatment. Anti-TF precipitates were boiled for 5 min to release the captured antigen from the antibody. Samples were cooled to 4 °C, and 20 U ml^{–1} peptide N-glycosidase F (Sigma) was added for 90 min at 37 °C and then boiled again for 5 min to inactivate the glycosidase. Treatment with O-glycosidase (25 mU ml^{–1}; Merck, Darmstadt) was performed for 3 h at 37 °C.

Immunoblotting. Tissue lysates or immunoprecipitates were separated by using a NuPAGE system with MOPS buffer and 10% BisTris gels. Proteins were transferred to poly(vinylidene difluoride) membranes (Invitrogen, Carlsbad). The membrane was blocked in 5% milk powder (in PBS/Tween) and incubated for 1.5 h in 5% milk powder containing the primary antibody (rabbit anti-mouse TF (2.5 µg ml^{–1}; American Diagnostica, Stamford) for immunoprecipitation, rabbit anti-mouse TF and rabbit anti-mouse phospho-TF (2 µg ml^{–1}) for specificity controls—see Supplementary Fig. 17, rabbit anti-integrin β₁ (dilution 1:1,000; Cell Signaling), rabbit anti-actin (dilution 1:200; Sigma), rabbit anti-phospho-Tie2 (dilution 1:250; R&D) and rabbit anti-Tie-2 antibody (dilution 1:250; Abcam)). Secondary goat anti-rabbit IgG (horseradish peroxidase-conjugated; Santa Cruz Biotechnology, Santa Cruz) was applied for 1 h. Alternatively, the membrane was incubated with horseradish peroxidase-conjugated concanavalin A (Sigma) to detect sugar moieties after immunoprecipitation. Then the membrane was first washed for 2 min with PBS and incubated overnight with the lectin solution (PBS containing Mg²⁺ and Ca²⁺). The next day, the blot was rinsed three times with PBS/Tween. Blots were developed with enhanced chemiluminescence solutions (Amersham Biosciences, Little Chalfont). For densitometric analysis of protein bands, the software Multi Gauge V3.0 (Fuji Film, Tokyo) was applied.

Cell-surface labelling and pull-down with N-hydroxysuccinimido-biotin. For amine-reactive biotinylation and isolation of cell surface proteins from isolated primary enterocytes, the Cell Surface Protein Isolation Kit (Pierce, Rockford) was used. Isolated proteins were separated on a 10% BisTris gel (Invitrogen), and TF antigen was analysed by immunoblotting.

Proximity ligation assay³³. Slides with adhering primary enterocytes were blocked and incubated with primary antibodies (monoclonal rat-anti-mouse TF (1H1)³⁴, 23.4 µg ml^{–1}, provided by Daniel Kirchhofer; rabbit polyclonal anti-integrin β₁, dilution 1:50; Cell Signaling Technology). Secondary antibodies (anti-rat and anti-rabbit) conjugated with unique DNA probes (Olink Bioscience, Uppsala) were added. Slides were evaluated with a Leica TCS SP5 confocal microscope. If TF and integrin β₁ antigens are closer than 30 nm, a fluorescence signal can be generated.

- Falcón, B. L. *et al.* Contrasting actions of selective inhibitors of angiopoietin-1 and angiopoietin-2 on the normalization of tumor blood vessels. *Am. J. Pathol.* **175**, 2159–2170 (2009).
- Stappenbeck, T. S. *et al.* Laser capture microdissection of mouse intestine: characterizing mRNA and protein expression, and profiling intermediary metabolism in specified cell populations. *Methods Enzymol.* **356**, 167–196 (2002).
- Petersen, L. C. *et al.* Characterization of recombinant murine factor VIIa and recombinant murine tissue factor: a human–murine species compatibility study. *Thromb. Res.* **116**, 75–85 (2005).
- Söderberg, O. *et al.* Direct observation of individual endogenous protein complexes *in situ* by proximity ligation. *Nature Methods* **3**, 995–1000 (2006).
- Kirchhofer, D., Moran, P., Bullens, S. & Peale, F. A monoclonal antibody that inhibits mouse tissue factor function. *J. Thromb. Haemost.* **3**, 1098–1099 (2005).

Structure and mechanism of a glutamate–GABA antiporter

Dan Ma^{1,3}, Peilong Lu^{1,3}, Chuangye Yan^{2,3}, Chao Fan¹, Ping Yin², Jiawei Wang² & Yigong Shi^{1,3}

Food-borne hemorrhagic *Escherichia coli*, exemplified by the strains O157:H7 and O104:H4 (refs 1, 2), require elaborate acid-resistance systems (ARs)³ to survive the extremely acidic environment such as the stomach (pH ≈ 2). AR2 expels intracellular protons through the decarboxylation of L-glutamate (Glu) in the cytoplasm and exchange of the reaction product γ -aminobutyric acid (GABA) with extracellular Glu. The latter process is mediated by the Glu–GABA antiporter GadC^{4,5}, a representative member of the amino-acid–polyamine–organocation superfamily of membrane transporters. The functional mechanism of GadC remains largely unknown. Here we show, with the use of an *in vitro* proteoliposome-based assay, that GadC transports GABA/Glu only under acidic conditions, with no detectable activity at pH values higher than 6.5. We determined the crystal structure of *E. coli* GadC at 3.1 Å resolution under basic conditions. GadC, comprising 12 transmembrane segments (TMs), exists in a closed state, with its carboxy-terminal domain serving as a plug to block an otherwise inward-open conformation. Structural and biochemical analyses reveal the essential transport residues, identify the transport path and suggest a conserved transport mechanism involving the rigid-body rotation of a helical bundle for GadC and other amino acid antiporters.

Other homologous amino-acid–polyamine–organocation (APC) family members include the key AR3 component arginine–agmatine (Arg–Agm) antiporter AdiC^{6,7} (Fig. 1a), the lysine–cadaverine antiporter CadB, and the putrescine–ornithine antiporter PotE. Structural analysis of AdiC^{8–11} revealed a conserved LeuT fold that is associated with the Na⁺-coupled symporters^{12–15} and a proton-coupled transporter ApcT¹⁶. Ligand-free AdiC exists in an outward-open conformation^{8,10}, and binding of Arg triggers a major structural rearrangement resulting in an occluded conformation⁹. Despite these advances, the transport mechanism for AdiC or any other amino acid antiporter remains poorly understood. The inward-open conformation has yet to be captured for any amino acid antiporter.

The full-length, wild-type (WT) GadC (residues 1–511), derived from the *E. coli* strain O157:H7, was purified to homogeneity. To characterize GadC, we reconstituted a proteoliposome-based transport assay (Fig. 1b), in which substrate transport was monitored through the detection of ³H-labelled Glu. Incorporation of GadC into the liposome allowed the rapid accumulation of Glu at pH 4.5 (Fig. 1b; Supplementary Fig. 1). Strikingly, substrate transport by GadC decreased sharply with increasing pH values; the accumulation of Glu after 15 s at pH 5.0, 5.5 and 6.0 was about 67%, 32% and 8% of that at pH 4.5 (Supplementary Fig. 1). At pH 6.5 and above, there was no detectable accumulation of Glu (Fig. 1b). Thus, substrate transport by GadC is strictly pH-dependent, with robust activity at pH 5.5 or below. This biochemical property not only helps in the survival of enterobacteria under acidic environment, but it may also be important for avoiding unnecessary proton efflux under neutral pH growth conditions.

AdiC has been thoroughly characterized, with elevated transport activities at pH 4 and 6 (ref. 17); however, AdiC still allows moderate

transport even under basic conditions such as pH 8 (ref. 17). We compared the activities of AdiC and GadC under identical sets of experimental conditions (Fig. 1c, d). The results confirmed and extended published findings. The total substrate accumulation in the proteoliposomes over 10 min at pH 9.0 was about 21% of that at pH 5.0 (Fig. 1c). In sharp contrast, GadC has no detectable activity at pH 6.5 or higher. Determination of the maximal transport activity (V_{\max}) and K_m for AdiC and GadC at different pH values confirmed these observations (Fig. 1d and Supplementary Fig. 2). Thus, in comparison with AdiC, GadC exhibits much more stringent dependence on pH for substrate transport.

Next we replaced GABA individually with 19 natural amino acids; the substrate transport for each amino acid was measured at pH 5.5 (Fig. 1e). In addition to Glu, GadC also efficiently transports three additional amino acids: Gln and, to a smaller extent, Met and Leu. This result was unexpected because, unlike Glu, none of the three amino acids carries a charge or can be protonated on the side chain. Notably, the V_{\max} for Gln is considerably larger than that for Glu or GABA at pH 5.5, whereas the K_m for Gln is comparable to that for GABA (Supplementary Fig. 3). In particular, GadC allows little, if any, transport of the amino acids Asp, Phe, Gly, His, Lys, Asn, Pro and Trp (Fig. 1e). Substrate transport for Asp and Asn was less than 5% of that for the chemically similar amino acids Glu and Gln, respectively. This analysis strongly indicates that GadC is highly selective in substrate transport.

Efflux of protons presumably contributes to a build-up of a negative potential within the lipid membrane, which is known to be unfavourable for continued substrate transport by amino acid antiporters such as AdiC¹⁷ and AspT¹⁸. We examined Glu–GABA exchange in the proteoliposome-based assay in the absence or presence of valinomycin (Supplementary Fig. 4). The results clearly demonstrate that positive potential within the proteoliposomes stimulated substrate transport, whereas negative potential led to decreased transport.

We crystallized the full-length GadC in the space group P2₁2₁2₁ at pH 8.0 and determined the structure by platinum-based single-wavelength anomalous dispersion (SAD) at 3.1 Å resolution (Supplementary Tables 1 and 2; Supplementary Figs 5 and 6). Each asymmetric unit contains two molecules of GadC, arranged in an antiparallel fashion (Supplementary Fig. 5). This packing pattern suggests that, similarly to AdiC¹⁰ and the transporter CIC¹⁹, the functional entity for GadC is likely to be a single molecule. GadC contains 12 TMs, with TM1 and TM6 each containing two short α -helices connected by a discontinuous stretch in the middle (Fig. 2a and Supplementary Fig. 7). The structure of GadC, together with the identification of two periplasmic loops in GadC²⁰, allows an unambiguous assignment of its membrane topology. In a similar manner to AdiC and other LeuT-type transporters, GadC contains two inverted repeats, TM1–TM5 and TM6–TM10, which are related to each other by a pseudo-two-fold axis (Supplementary Fig. 8a). The periplasmic and cytoplasmic sides of GadC are highly charged (Supplementary Fig. 8b).

¹Ministry of Education Protein Science Laboratory, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China. ²State Key Laboratory of Biomembrane and Membrane Biotechnology, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China. ³Tsinghua–Peking Joint Center for Life Sciences, Center for Structural Biology, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China.

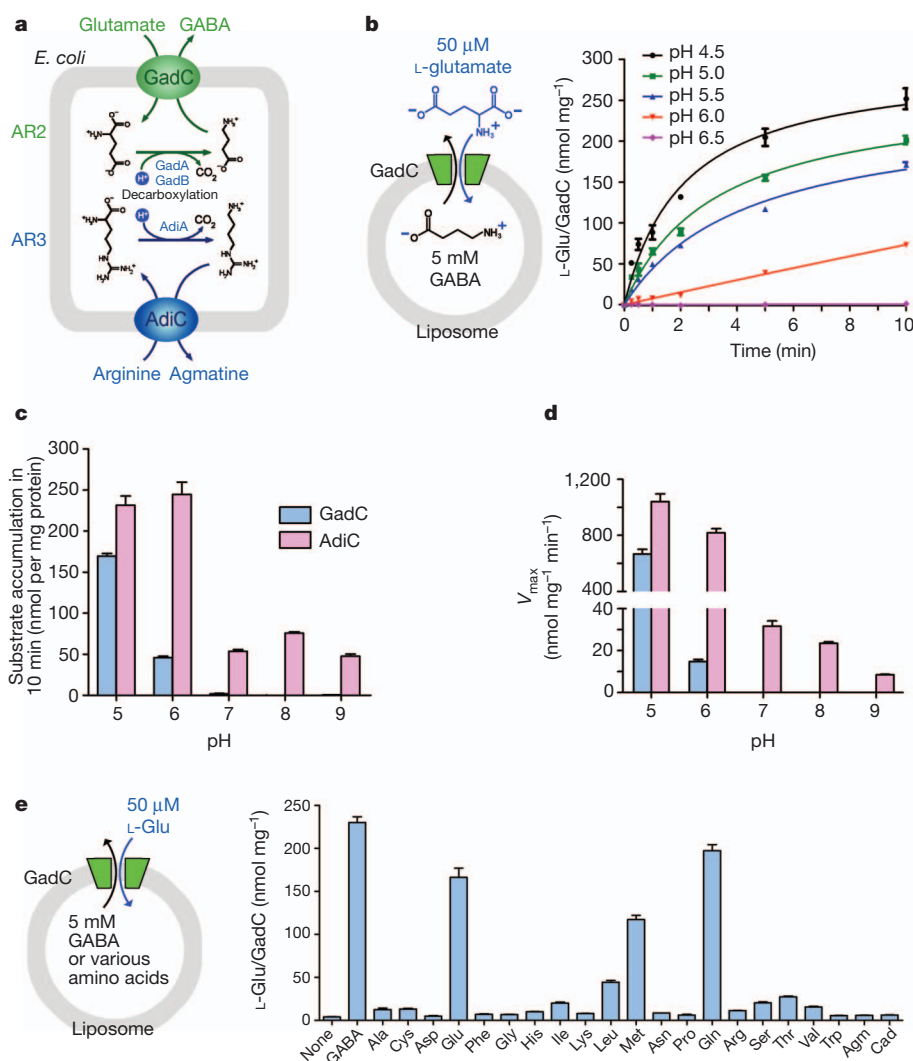


Figure 1 | Functional characterization of GadC. **a**, Schematic diagram of AR2 and AR3 in *E. coli*. The antiporter GadC of AR2 exchanges extracellular Glu for intracellular GABA, resulting in the net efflux of one proton per cycle. Glu is decarboxylated by GadA/B to become GABA in cells. The Arg–Agm antiporter AdiC and the decarboxylase AdiA are the equivalents of GadC and GadA/B, respectively. **b**, GadC shows pH-dependent transport in the proteoliposome assay. ³H-labelled Glu was present at roughly 0.19 μ M to permit the measurement of transport. The transport activity is robust at pH 4.5 and rapidly decreases with increasing pH, with no detectable transport activity at pH 6.5 or above. **c**, Comparison of pH-dependent substrate transport of AdiC and GadC. GadC exhibits stringent pH dependence, with no detectable transport activity at pH values higher than 7.0. By contrast, AdiC has considerable transport activity at pH 7.0, 8.0 and 9.0. **d**, Comparison of V_{max} for AdiC and GadC at different pH values. **e**, The substrate specificity of GadC-mediated transport in the proteoliposome assay. Substrate transport was measured at pH 5.5 for 10 min. GadC only allows transport of GABA, Glu, Gln and (to a smaller extent) Met and Leu. All error bars represent the s.d. for three independent experiments.

In comparison with AdiC and other homologues, GadC contains a unique, extended C-terminal fragment (Supplementary Fig. 9). In contrast with reported structures of AdiC^{8–11}, GadC seems to adopt an inward-open conformation (Fig. 2a). The open path leads to a negatively charged environment (Fig. 2b), where substrate-binding residues are likely to be located. However, the C-terminal fragment (residues 477–511) forms a folded domain and completely blocks the path to the putative substrate-binding site. The C-terminal fragment, with clear electron density (Supplementary Fig. 10a), is hereafter referred to as the C-plug. The observation of a blocked transport path in GadC is consistent with the fact that the crystals were generated at pH 8.0, at which no transport activity could be detected (Fig. 1c).

The location of the C-plug strongly suggests that its displacement is a prerequisite for the transport activity of GadC. To examine this, we generated a C-plug-deleted GadC variant (residues 1–470) and measured its ability to permeate Glu–GABA. Strikingly, whereas WT GadC showed little substrate transport over 60 min at pH 6.5, GadC (1–470) showed a significant level of Glu accumulation (Fig. 3a).

Additional measurements show that the transport activity of GadC (1–470) was rapidly decreased with increasing pH and became undetectable at pH 7.5 or above (Fig. 3b). Overall, deletion of the C-plug in GadC shifted its pH-dependent substrate transport towards a higher pH.

The C-plug contains several basic amino acids and makes multiple intra-domain and inter-domain hydrogen bonds (Fig. 3c and Supplementary Fig. 10b). The tightly folded conformation of the C-plug is stabilized by two centrally located basic residues, His 491 and Arg 499. His 491 donates a hydrogen bond to Ser 500, whereas Arg 499 makes three hydrogen bonds to the main-chain atoms of Phe 492 and Leu 494. At one end of the plug, Tyr 503 forms a hydrogen bond to Ala 487. These intra-domain interactions are complemented by inter-domain contacts. At the other end of the plug, the guanidinium group of Arg 497 donates a hydrogen bond to Gln 98 on TM3. In addition, His 502 makes a hydrogen bond to Arg 314 on TM8, whereas the main chain atoms of Val 477 and Ser 484 form hydrogen bonds with Gln 321 on TM8 and Glu 226 on TM6.

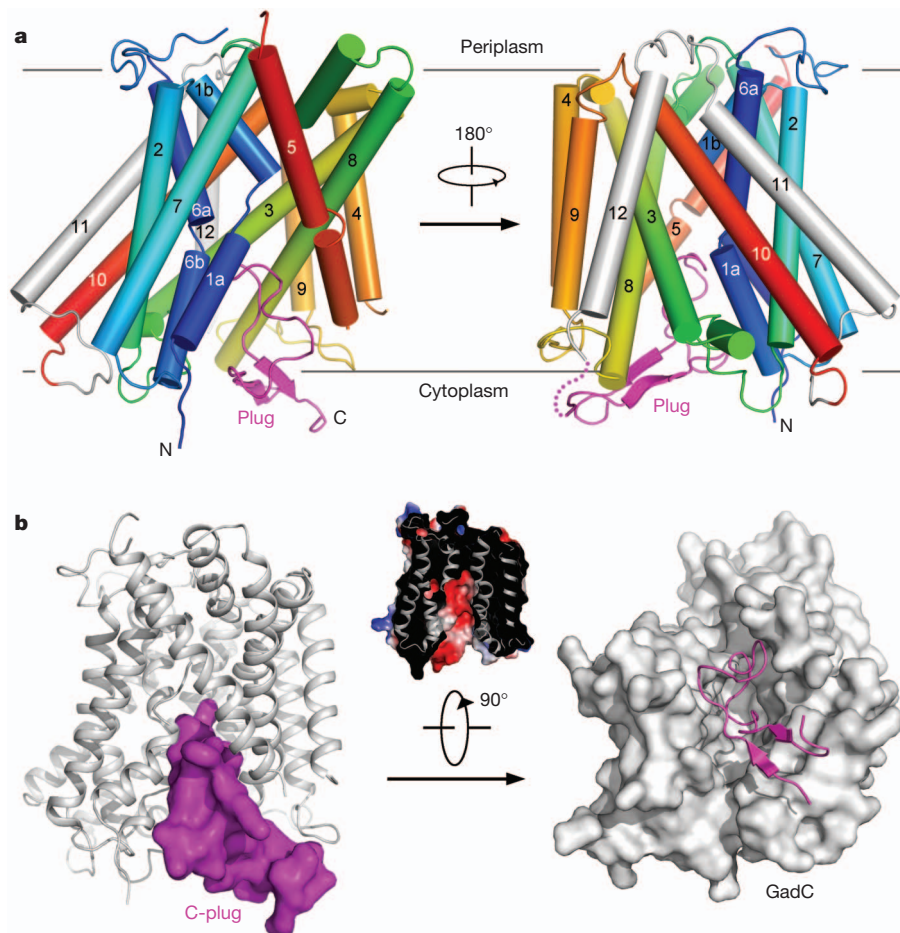


Figure 2 | Overall structure of GadC. **a**, Overall structure of the WT full-length GadC. TM1–TM10 are rainbow-coloured, with TM1 in blue and TM10 in red. TM11 and TM12 are shown in grey. The C-terminal fragment (C-plug) is coloured magenta. **b**, The C-plug of GadC blocks an otherwise inward-open

Perturbation of these interactions is predicted to compromise the stability of the C-plug and consequently to facilitate its displacement. To examine this prediction, we generated five GadC variants—H491A,

conformation. Two perpendicular views of GadC are shown. The C-plug blocks access to the negatively charged substrate-binding cleft (inset). All structural figures were prepared with PyMol³⁰.

R497A, R499A, H502A and Y503A—each carrying a missense mutation on a critical residue in the C-plug. We individually investigated substrate transport by these GadC variants at pH 6.5. Whereas the WT GadC exhibited little activity, all five GadC variants allowed varying but significant levels of substrate transport (Fig. 3d). In particular, the transport activities for GadC-H491A and GadC-R497A were similar to that of GadC (1–470) (Fig. 3d). By contrast, two additional GadC variants, H495A and H511A, allowed considerably less Glu accumulation, consistent with the observation that these mutations do not affect any critical interactions in the C-plug.

The closed conformation of GadC is attained not only by the C-plug in the cytoplasm but also by the L7 loop at the periplasmic side (Supplementary Fig. 11a). The L7 loop interacts with surrounding structural elements through a combination of hydrogen bonds and van der Waals contacts (Supplementary Fig. 11b, c). We speculate that,

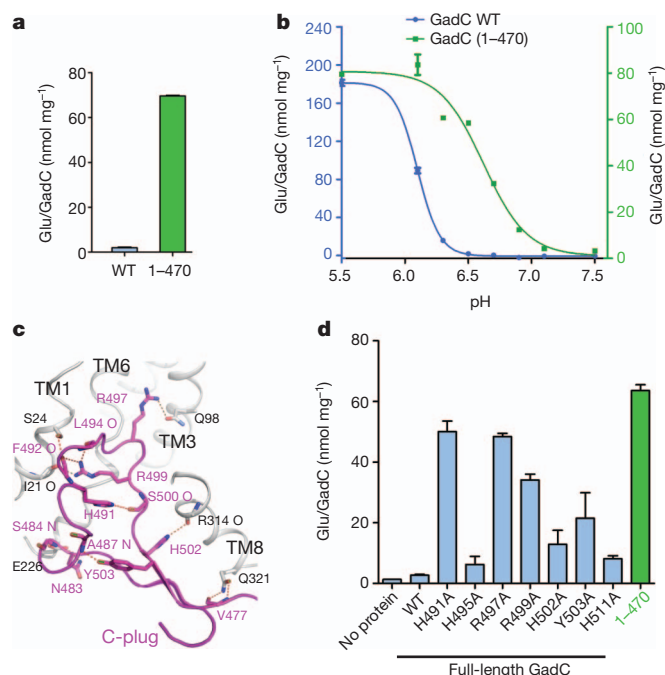


Figure 3 | The C-plug regulates substrate transport. **a**, Truncation of the C-plug (residues 471–511) allowed transport of Glu at pH 6.5. Shown here is the total accumulation of Glu in the proteoliposomes in 60 min. **b**, Deletion of the C-plug shifts the pH-dependent transport activity of GadC towards higher pH values. **c**, The C-plug interacts with surrounding structural elements through multiple hydrogen bonds. The C-plug is coloured magenta and surrounding structural elements are shown in grey. **d**, Structural integrity of the C-plug is important for the transport activity of the WT GadC. Although the WT GadC does not allow apparent transport of substrates at pH 6.5, several missense mutants acquired this ability. These mutations probably led to compromised interactions between the C-plug and its surrounding structural elements. All error bars represent the s.d. for three independent experiments.

during each cycle of transport, the L7 loop must be displaced, at least transiently, to allow the passage of substrate molecules. The C α –C α distance between residue 267 on L7 and residue 364 on TM10 is about 5.6 Å (Supplementary Fig. 11d), which is ideal for disulphide bond formation if these two residues are replaced by Cys. We generated a double mutation GadC-L267C/N364C and subjected the purified protein to oxidation by *o*-phenanthroline copper complex²¹. The oxidized GadC variant (L267C/N364C) showed undetectable substrate transport in the absence of the reducing agent dithiothreitol but restored substrate transport in the presence of dithiothreitol (Supplementary Fig. 11d). By contrast, the oxidized WT GadC showed similar levels of substrate transport in the absence or presence of dithiothreitol.

The available structures of AdiC^{8–11} greatly facilitated the identification of gating residues in GadC. The substrate transport path is sandwiched axially between the C-plug and the L7 loop and surrounded laterally by TM1, TM3, TM6, TM8 and TM10 (Fig. 4a). Sequence alignment with AdiC and other amino acid antiporters identified six potential gating residues in GadC: Tyr 96, Tyr 214, Glu 218, Trp 308, Tyr 378 and Tyr 382 (Supplementary Fig. 9). These residues are located within or in close proximity to the putative transport path (Fig. 4a). Comparison of the spatial locations for these amino acids revealed a pattern that differed from that in AdiC of the outward-open conformation⁸ or Arg-bound occluded conformation⁹ (Supplementary Fig. 12).

The distal gate in AdiC comprises three amino acids, with Glu 208 hydrogen bonded to Tyr 93 and Tyr 365 (ref. 9). Glu 208 and Tyr 365 in AdiC correspond to Glu 218 and Tyr 382 in GadC, respectively (Supplementary Fig. 9). Tyr 378 is also located within hydrogen-bonding distance of Glu 218. Tyr 96 in GadC, which may correspond to Tyr 93 in AdiC, is located roughly 5 Å away from the position of Tyr 93

(Supplementary Fig. 12). The middle gate residue Trp 293 in AdiC corresponds to Trp 308 in GadC, which is farther from the axial centre of transport, probably reflecting the nature of the inward-open conformation. The proximal gate residue, Trp 202, in AdiC closes on binding to Arg and occludes the substrate molecule from the periplasm⁹; the corresponding residue remains to be conclusively identified in GadC. The nearby residues Phe 210 and Tyr 214 are located in different positions from that of Trp 202 in AdiC. This analysis suggests major conformational changes for these putative gating residues during the substrate transport cycle. Several other amino acids are located within the putative transport path and may have a function in transport. In particular, Tyr 30 on TM1 donates a hydrogen bond to Glu 304 on TM8 (Supplementary Fig. 13a).

To corroborate the structural analysis, we generated six GadC variants—Y30A, E218A, E304A, W308A, Y378A and Y382A—each with a missense mutation targeting a putative gating residue. In contrast with WT GadC, each mutation caused at least 90% decrease in substrate transport (Supplementary Fig. 13b). GadC-W308A, which affects the putative middle gate residue Trp 308, only retained about 2% of the WT activity. By contrast, the GadC variant M25A, which does not affect any putative transport residue, showed about 74% of the WT transport activity.

Under the assumption of a conserved transport mechanism between GadC and AdiC, the inward-open conformation of GadC, with omission of the C-plug, probably reflects a distinct state of AdiC during substrate transport. Structural alignment between outward-open AdiC⁸ and GadC reveals a striking pattern of conformational changes that are concentrated in a helical bundle comprising TM1, TM2, TM6 and TM7 (Fig. 4b). For simplicity of discussion, these four TMs are collectively named the gate domain, and the rest is referred to

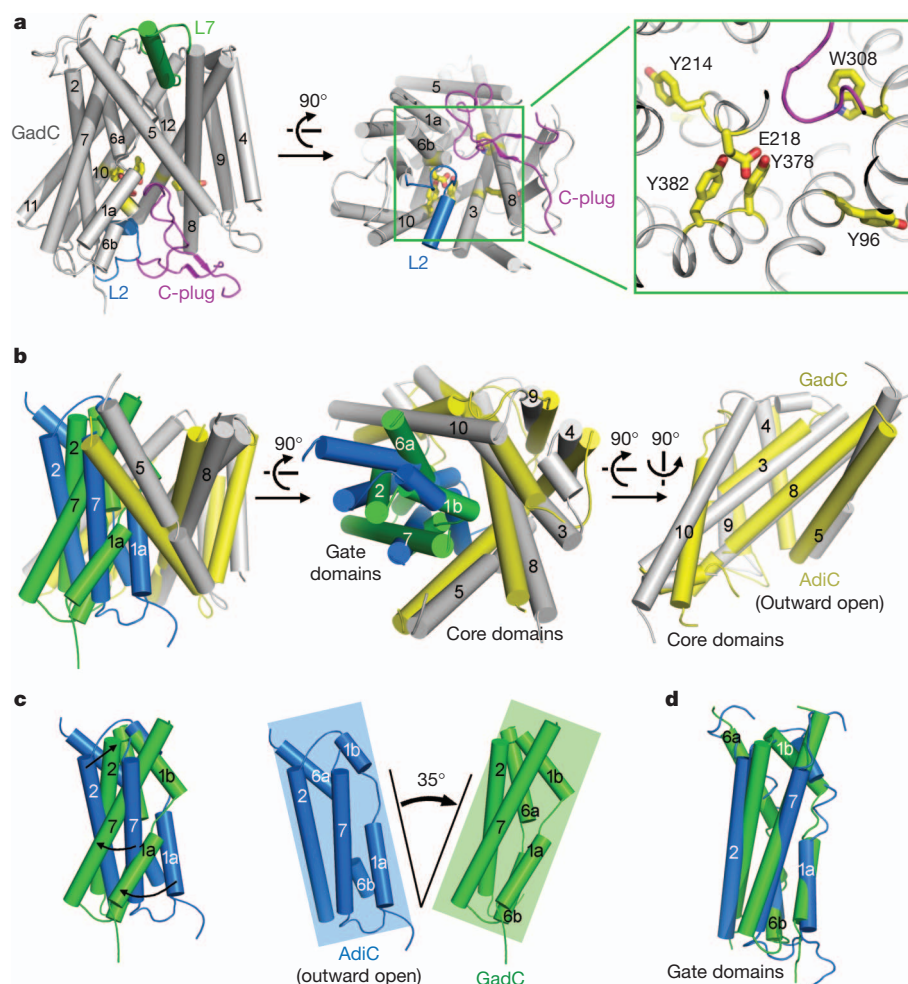


Figure 4 | The transport path in GadC.

a, Identification of key amino acids in the transport path. Sequence alignment and structural analysis identify six amino acids that may be essential in substrate transport. These residues are shown in a close-up view in the right panel. **b**, The gate domain in GadC, comprising TM1, TM2, TM6 and TM7, undergoes the most pronounced conformational changes compared with those [of the outward-open AdiC⁸. The core domain showed only minor changes. The gate and core domains of GadC are coloured green and grey, respectively; the gate and core domains of AdiC are coloured blue and yellow, respectively. **c**, The conformational changes of the gate domain amount to a rigid-body rotation of about 35°. **d**, Structural overlay of the gate domains from GadC (green) and AdiC (blue).

as the core domain. In contrast with that in AdiC, the gate domain in GadC seems to rotate clockwise by about 35° (Fig. 4c), resulting in an outward-closed and inward-open conformation in GadC. Structural alignment between the isolated gate domains of AdiC and GadC revealed only minor changes in these four TMs (Fig. 4d), suggesting a rigid-body movement between the gate and the core domains. This structural analysis supports the notion of alternating access for membrane transporters^{22,23}. The four TMs in the gate domain undergo the most drastic structural rearrangement among all 12 TMs of GadC (Supplementary Fig. 14). In comparison with the gate domain, TMs of the core domain have a considerably smaller degree of conformational changes, particularly TM3, TM8 and TM9.

The mechanistic conservation may go beyond the amino acid antiporters. Mhp1 was recently found to switch from outward-open to inward-open conformation through a rigid-body rotation involving two helical bundles²⁴. After the exit of Na⁺, the Na⁺-galactose transporter vSGLT was also thought to undergo a minor rigid-body movement involving two helical bundles²⁵. Of the two helical bundles, one contains TM3, TM4, TM8 and TM9, and the other comprises TM2, TM6 and TM7; this is true for Mhp1 (ref. 24), vSGLT²⁵ and AdiC/GadC, for all of which at least two distinct conformations of a transport cycle have been structurally characterized. The helical bundle that executes the rigid-body movement is proposed to be the gate domain (TM1, TM2, TM6 and TM7) in GadC and the hash motif (TM3, TM4, TM8 and TM9) in Mhp1 (ref. 24). Thus, the moving and non-moving portions of GadC are exactly the reciprocal of those in Mhp1. The choice of the core domain in GadC is justified by the following analysis. Superposition between AdiC and GadC yields a root mean squared deviation of 2.98 Å over 167 aligned Cα atoms if TM3, TM4, TM5, TM8, TM9 and TM10 are treated as the non-moving portion, and about 3.95 Å over 143 aligned Cα atoms if TM1, TM2, TM5, TM6, TM7 and TM10 are treated as the non-moving portion.

Structural elucidation of the Glu-GABA antiporter GadC is a step towards a detailed, mechanistic understanding of the amino acid antiporters. Many questions remain (see Supplementary Discussion). At present we have little information about how and to what extent the C-plug of GadC is dislodged during substrate transport. We have yet to explain the pH-dependent transport activity by amino acid antiporters such as GadC and AdiC. Conclusive answers to these questions require additional biochemical and structural investigation.

METHODS SUMMARY

All constructs were generated by standard PCR-based protocol. GadC was over-expressed in *E. coli* BL21(DE3). The proteins were purified to homogeneity by affinity chromatography and gel filtration. Crystals of GadC were grown by the hanging-drop vapour-diffusion method. The crystals belong to the space group *P*₂₁₂₁₂; they were flash-frozen in a cold nitrogen stream at 100 K. X-ray data were collected at the Shanghai Synchrotron Radiation Facility (SSRF) beamline BL17U and SPring-8 beamline BL41XU. Data were processed with HKL-2000 (ref. 26). The platinum positions were determined with the program SHELXD²⁷. Cross-crystal averaging with all three data sets combined with solvent flattening, histogram matching and non-crystallographic symmetry (NCS) averaging in DMMulti gave a map of sufficient quality for model building. An initial model was built into the experimental map by using COOT²⁸. The sequence docking was aided with the selenium sites in the anomalous difference Fourier map. The structure was refined with PHENIX²⁹. Proteoliposome assays were performed to determine transport activities of various GadC mutants.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 October 2011; accepted 3 February 2012.

Published online 11 March 2012.

- Riley, L. W. *et al.* Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**, 681–685 (1983).
- Rohde, H. *et al.* Open-source genomic analysis of Shiga-toxin-producing *E. coli* O104:H4. *N. Engl. J. Med.* **365**, 718–724 (2011).
- Foster, J. W. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nature Rev. Microbiol.* **2**, 898–907 (2004).

- Hersh, B. M., Farooq, F. T., Barstad, D. N., Blankenhorn, D. L. & Slonczewski, J. L. A glutamate-dependent acid resistance gene in *Escherichia coli*. *J. Bacteriol.* **178**, 3978–3981 (1996).
- Castanie-Cornet, M. P., Penfound, T. A., Smith, D., Elliott, J. F. & Foster, J. W. Control of acid resistance in *Escherichia coli*. *J. Bacteriol.* **181**, 3525–3535 (1999).
- Iyer, R., Williams, C. & Miller, C. Arginine- α -glutamate antiporter in extreme acid resistance in *Escherichia coli*. *J. Bacteriol.* **185**, 6556–6561 (2003).
- Gong, S., Richard, H. & Foster, J. W. YjdE (AdiC) is the arginine- α -glutamate antiporter essential for arginine-dependent acid resistance in *Escherichia coli*. *J. Bacteriol.* **185**, 4402–4409 (2003).
- Gao, X. *et al.* Structure and mechanism of an amino acid antiporter. *Science* **324**, 1565–1568 (2009).
- Gao, X. *et al.* Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463**, 828–832 (2010).
- Fang, Y. *et al.* Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* **460**, 1040–1043 (2009).
- Kowalczyk, L. *et al.* Molecular basis of substrate-induced permeation by an amino acid antiporter. *Proc. Natl Acad. Sci. USA* **108**, 3935–3940 (2011).
- Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na⁺/Cl[−]-dependent neurotransmitter transporters. *Nature* **437**, 215–223 (2005).
- Ressl, S., Terwisscha van Scheltinga, A. C., Vonnrhein, C., Ott, V. & Ziegler, C. Molecular basis of transport and regulation in the Na⁺/betaine symporter BetP. *Nature* **458**, 47–52 (2009).
- Faham, S. *et al.* The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science* **321**, 810–814 (2008).
- Weyand, S. *et al.* Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322**, 709–713 (2008).
- Shaffer, P. L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. Structure and mechanism of a Na⁺-independent amino acid transporter. *Science* **325**, 1010–1014 (2009).
- Fang, Y., Kolmakova-Partensky, L. & Miller, C. A bacterial arginine- α -glutamate exchange transporter involved in extreme acid resistance. *J. Biol. Chem.* **282**, 176–182 (2007).
- Sasahara, A., Nanatani, K., Enomoto, M., Kuwahara, S. & Abe, K. Substrate specificity of the aspartate:alanine antiporter (AspT) of *Tetragenococcus halophilus* in reconstituted liposomes. *J. Biol. Chem.* **286**, 29044–29052 (2011).
- Robertson, J. L., Kolmakova-Partensky, L. & Miller, C. Design, function and structure of a monomeric ClC transporter. *Nature* **468**, 844–847 (2010).
- Waterman, S. R. & Small, P. L. Identification of sigma S-dependent genes associated with the stationary-phase acid-resistance phenotype of *Shigella flexneri*. *Mol. Microbiol.* **21**, 925–940 (1996).
- Kobashi, K. Catalytic oxidation of sulfhydryl groups by o-phenanthroline copper complex. *Biochim. Biophys. Acta* **158**, 239–245 (1968).
- Abramson, J. *et al.* The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport. *FEBS Lett.* **555**, 96–101 (2003).
- Jardetzky, O. Simple allosteric model for membrane pumps. *Nature* **211**, 969–970 (1966).
- Shimamura, T. *et al.* Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328**, 470–473 (2010).
- Watanabe, A. *et al.* The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature* **468**, 988–991 (2010).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779 (2002).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
- DeLano, W. L. The PyMOL Molecular Graphics System, (<http://www.pymol.org>) (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Huang and J. He (at SSRF beamline BL17U) and N. Shimizu, T. Kumasaka and S. Baba at the Spring-8 beamline BL41XU for on-site assistance, and N. Yan for discussion and comments on the manuscript. This work was supported by funds from the Ministry of Science and Technology (grant 2009CB918801), National Natural Science Foundation, and Beijing Municipal Commissions of Education and Science and Technology.

Author Contributions D.M., P.L. and Y.S. designed all experiments. D.M., P.L., C.Y., C.F. and P.Y. performed the experiments. D.M., P.L., C.Y., C.F., P.Y., J.W. and Y.S. analysed the data. D.M., P.L., C.Y., J.W. and Y.S. contributed to manuscript preparation. Y.S. wrote the manuscript.

Author Information The atomic coordinates and the structure factor file have been deposited in the Protein Data Bank under accession numbers 4DJK and 4DJI. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.S. (shi-lab@tsinghua.edu.cn).

METHODS

Protein preparation. The cDNA of full-length GadC was subcloned into pET15b (Novagen). Overexpression of GadC was induced by 0.2 mM isopropyl β -D-thiogalactoside (IPTG) when the cell density reached a D_{600} of 1.2. After growth at 37 °C for 4 h, the cells were collected, homogenized in buffer containing 25 mM Tris-HCl pH 8.0 and 150 mM NaCl. After further disruption with a French press, cell debris was removed by low-speed centrifugation for 10 min. The supernatant was collected and ultracentrifuged for 1 h at 150,000g. The membrane fraction was collected and homogenized with buffer containing 25 mM Tris-HCl pH 8.0 and 150 mM NaCl. *N*-Octyl- β -D-glucopyranoside (β -OG; Anatrace) was added to the membrane suspension to a final concentration of 2% (w/v) and then incubated for 2 h at 4 °C. After another ultracentrifugation step at 150,000g for 30 min, the supernatant was collected and loaded on Ni^{2+} -nitrilotriacetate affinity resin (Ni-NTA; Qiagen), followed by a wash with 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 0.4% *n*-nonyl- β -D-maltopyranoside (NM; Anatrace). After proteolytic removal of the hexahistidine (His_6) tag on the column, GadC was eluted with buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.4% NM. After concentration to 10–15 mg ml⁻¹, GadC was further purified by gel filtration (Superdex-200 10/30; GE Healthcare). The buffer for gel filtration contained 25 mM Tris-HCl pH 8.0, 150 mM NaCl and various detergents. The peak fractions were collected. The GadC mutants were generated with a standard PCR-based strategy and were subcloned, over-expressed and purified in the same way as the WT protein.

Crystallization. The hanging-drop vapour-diffusion method was performed at 18 °C during crystallization. Crystals belonging to the space group $P2_12_12_1$ were obtained with protein purified in the presence of 0.2% *n*-nonyl- β -D-glucopyranoside (β -NG; Anatrace) and 0.023% *n*-dodecyl-*N,N*-dimethylamine-*N*-oxide (LDAO; Anatrace). The crystallization buffer was 21% PEG400, 100 mM Tris-HCl pH 8.0, 100 mM NaCl, and 325 mM sodium acetate. Rod-shaped crystals appeared overnight and typically grew to full size in about 1 week. Crystals were dehydrated by exposing the drops to air for 5 min. The best diffraction reached 2.95 Å at SSRF beamline BL17U. Platinum derivatives were obtained by soaking the crystals for 48 h in mother liquor containing 10 mg ml⁻¹ $\text{K}_2\text{Pt}(\text{NO}_2)_4$. Seleno-L-methionine-incorporated crystals were also obtained and reach similar diffraction with heavy-atom-derived crystals. Diffraction data for heavy-atom and selenomethionine derivatives were collected at SPring-8 beamline BL41XU.

Data collection and structure determination. All anomalous diffraction data, including Pt-SAD and SeMet-SAD data, were collected at SPring-8 beamline BL41XU and processed with the package HKL-2000 (ref. 26) with routine procedures. The diffraction images from the severely anisotropic native crystal were collected at SSRF beamline BL17U and integrated with DENZO²⁶. Before the .x files were inputted into SCALEPACK²⁶ for merging and scaling, the anisotropic ellipsoidal truncations on the .x files were performed with the special version of the ellipsoidal truncation program provided by the University of California, Los Angeles, MBI Diffraction Anisotropy Server³¹. The applied resolution limits along the a^* , b^* and c^* directions are 3.77, 3.31 and 3.10 Å, respectively, on the basis of on the criterion of F/σ larger than 3.0. The pruned data were then used for structural determination and refinement. Further processing was carried out with programs from the CCP4 suites³². Data collection statistics are summarized in Supplementary Table 1, and Supplementary Table 2 compares the data completeness before and after the anisotropic truncation.

The platinum positions were determined with the program SHELXD³³. The identified platinum sites were then refined, and initial phases were calculated in the program PHASER³⁴. Cross-crystal averaging with all the three data sets combined with solvent flattening, histogram matching and non-crystallographic symmetry (NCS) averaging in DMMulti³⁵ gave a map of sufficient quality for model building. An initial model was built into the experimental map with COOT²⁸. The sequence docking was aided with the selenium sites in the anomalous difference Fourier map. The structure was refined with PHENIX²⁹.

Preparation of oxidized protein. Oxidation of the WT GadC and the GadC mutant L267C/N364C was performed with *o*-phenanthroline copper complex²¹. The oxidation system comprised 25 mM MES buffer pH 5.5, 150 mM KCl, 2.5 mg ml⁻¹ protein and 0.9 mM *o*-phenanthroline copper complex. The reaction was performed for 2 h on ice.

Preparation of liposomes and proteoliposomes. Liposomes of *E. coli* polar lipid (Avanti) was prepared using a standard protocol as described previously³⁶. For the study of membrane potential on substrate transport by GadC, liposomes were loaded with 5 mM GABA and either 120 mM sodium phosphate pH 5.5 or potassium phosphate pH 5.5. For all other transport assays of GadC, liposomes were loaded with various choices of buffer system (25 mM) depending on the assay purposes, 150 mM KCl and 5 mM GABA (or other amino acid and their derivatives). For AdiC, 5 mM GABA was replaced with 5 mM agmatine or arginine. WT or mutant GadC or WT AdiC were incorporated with liposomes to form proteoliposomes by incubation with pre-extruded liposomes together with 1.25% β -OG (Anatrace) at a concentration of 5 $\mu\text{g mg}^{-1}$ lipids. β -OG was removed by incubation overnight with 400 mg ml⁻¹ Bio-Beads SM2 (Bio-Rad). The proteoliposomes were harvested by ultracentrifugation for 1 h at 150,000g and rinsed twice with resuspension buffer (various choices of buffer system (25 mM), 150 mM KCl). The proteoliposomes were resuspended with the resuspension buffer to a final lipid concentration of 100 mg ml⁻¹.

In vitro transport assay. All transport assays were performed at 25 °C. For the study of membrane potential on substrate transport by GadC, the reaction was initiated by adding proteoliposomes (2 μl) to 100 μl of external buffer containing 120 mM sodium phosphate pH 5.5 or potassium phosphate pH 5.5, 50 μM unlabelled L-glutamic acid and 1 μCi of L-[³H]glutamic acid (specific radioactivity 51.1 Ci mmol⁻¹; PerkinElmer Life Sciences), with or without 1 $\mu\text{g ml}^{-1}$ valinomycin. For all other L-glutamic acid uptake assays of GadC, the reaction was initiated by adding proteoliposomes (2 μl) to 100 μl of external buffer containing 25 mM pH buffer, 150 mM KCl, 50 μM unlabelled L-glutamic acid and 1 μCi of L-[³H]glutamic acid. The final concentration of L-[³H]glutamic acid in the external buffer was about 0.19 μM . For AdiC, unlabelled and ³H-labelled L-glutamic acid were replaced by unlabelled and ³H-labelled L-arginine. The uptake of ³H-labelled substrate was stopped at the indicated time points by rapidly filtering the reaction solution through a 0.22- μm GSTF filter (Millipore) and washed with 2 ml of ice-cold wash buffer (25 mM glycine pH 9.5, 150 mM KCl). The filter was then taken for liquid scintillation counting. All experiments were repeated at least three times. The reactions lasted for various lengths of time depending on the different assay purposes.

For determination of V_{max} and K_m , the same substrates were used on both sides of proteoliposomes (Glu/Glu, Gln/Gln, GABA/GABA for GadC; Arg/Arg for AdiC). The chosen time points were within the linear range of substrate accumulation. The preparation of proteoliposomes and the transport assay process were as described above.

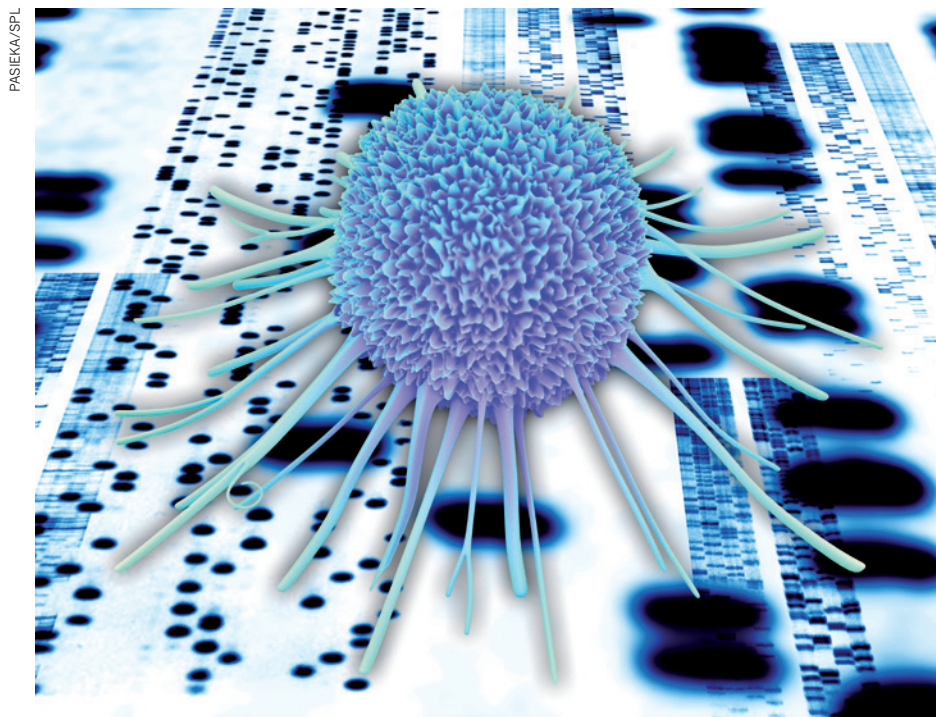
- Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).
- Bailey, S. The Ccp4 suite—programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
- Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779 (2002).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
- Cowan, K. dm: an automated procedure for phase improvement by density modification. *CCP4 ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–38 (1994).
- Veenhoff, L. M. & Poolman, B. Substrate recognition at the cytoplasmic and extracellular binding site of the lactose transport protein of *Streptococcus thermophilus*. *J. Biol. Chem.* **274**, 33244–33250 (1999).

CAREERS

UNITED STATES State ruling bars graduate students from forming a union **p.638**

COLUMN Embracing different ideas needs more than a change of scenery **p.639**

NATUREJOBS For the latest career listings and advice www.naturejobs.com



Computer reconstruction of a cancer cell on a DNA autoradiogram.

EPIGENETICS

Marked for success

The growing field of cancer epigenetics demands computational expertise and translational research experience. Qualified practitioners are in high demand.

BY HEIDI LEDFORD

When Constellation Pharmaceuticals first called to recruit venture capitalist Mark Goldsmith to be its chief executive in 2009, he was sceptical. Although Goldsmith was looking to change careers, he worried that the young biopharmaceutical company was heading into murky waters. The firm in Cambridge, Massachusetts, was focusing on epigenetics — the study of heritable changes in gene expression that are not due to changes

in DNA sequence. It planned to create cancer treatments that correct the abnormal patterns of epigenetic DNA modifications seen in tumours. “I took some convincing,” he says. “This was not an easy class of targets to go after; they were all unprecedented targets with incompletely understood biology.”

Despite his qualms, Goldsmith took the helm and, nearly three years later, epigenetics has become a hot topic in oncology drug discovery. In January, biotechnology giant Genentech in South San Francisco, California, added its

own vote of confidence to the field by investing US\$95 million in a partnership with Constellation, which is now hoping to add another 10 scientists to its current roster of 70.

Epigenetics, and cancer epigenetics in particular, is a bright spot in an otherwise stark biomedical-research funding and jobs landscape. Those with the right skills and background — computational and bioinformatics training, familiarity with and interest in translational research, and an intimate knowledge of molecular biology and cancer research techniques — have plenty of opportunities from which to choose. In particular, computational skills are so sought after that they alone could be a bridge to the sub-discipline. “It’s a really hot field,” says Benjamin Garcia, a chemist at Princeton University in New Jersey. “I wouldn’t be surprised if in five to ten years, you’re going to see a lot of universities with epigenetics departments.”

COMING OF AGE

Genetic mutations are not the only way to alter gene expression and protein function. Methyl groups added to DNA can silence a gene, as can chemical changes made to proteins called histones, which package the DNA in chromosomes. The modifications are exquisitely complex: the effect of one epigenetic change can be shaped by other modifications found nearby, and the epigenetic state of a cell will vary depending on the cell’s identity and maturity.

By the 1990s, researchers knew that the epigenetic state of a cancer cell was often in disarray. DNA methylation, for example, was markedly reduced in some tumours, unleashing gene-expression programs that were normally kept under lock and key. “The cancer genome was grossly different,” says Susan Clark, an epigeneticist at the University of New South Wales in Australia. “It was an amazing discovery.”

But many in the field needed further convincing before accepting that this epigenetic ‘chaos’ promoted changes in gene expression and, ultimately, led to cancer. That scepticism was not limited to industry; academics also worried that a career in the field would bring funding struggles and rejections from high-impact journals. “Clearly it was a risk, even ten years ago, to somebody’s career to dedicate themselves to an area that seemed to have a lot of hand-waving,” says Clark. “Now that’s changed; it is certainly a growth area for young scientists.”

As mounting evidence pointed to the importance of epigenetic changes in cancer, government funders began making significant investments in the field. The US National

DIVERSITY

PhD completion rates

In the hope of boosting degree completion rates, the US Council of Graduate Schools (CGS) in Washington DC is to examine attrition of minority students in science, technology, engineering and maths (STEM) programmes. The CGS will analyse data from 21 public and private universities for those entering programmes between 1992 and 2012. It aims to visit sites and interview students, faculty members and administrators to identify impediments to completion, and develop tools to remove them. Previous studies found that completion rates of minority students for STEM PhDs were significantly lower than those of non-minority students, notes Robert Sowell, vice-president for programmes and operations at CGS.

UNITED STATES

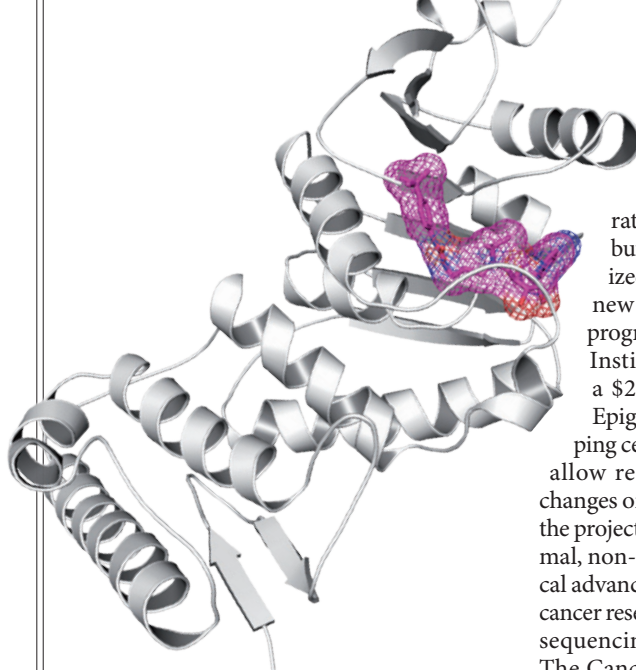
Unions banned

Michigan Governor Rick Snyder has banned graduate-student research assistants in public universities from unionizing following the efforts of 1,200 students to organize a union in April 2011. Snyder said in a statement that research assistants are students and giving them public-employee status and union representation would alter the student-teacher relationship. This is the latest action against US graduate-student unions. In 2004, New York University's union was disbanded under a state labour-board decision. Student representatives from Michigan State University in East Lansing and University of Michigan in Ann Arbor did not respond to interview requests.

PARTNERSHIPS

Postdoc opportunities

The California Institute for Quantitative Biosciences (QB3), part of the University of California, will hire up to 15 postdocs in a collaboration with drug firm Pfizer that expands a 2009 agreement to discover and develop technologies and drugs. Postdocs will be funded for two years in areas such as cardiovascular disease, immunology, neuroscience and oncology. They will learn to work with industry, says QB3 director Regis Kelly, who notes that this is a key activity given that many will go on to seek industry positions. Pfizer contributed US\$9.5 million to the original partnership and will provide at least the same level of funding again, says Ron Newbold, Pfizer's vice-president for strategic research partnerships.



An example of one of Epizyme's inhibitors interacting with an epigenetic enzyme.

► Cancer Institute (NCI) in Bethesda, Maryland, has several programmes dedicated to epigenetics, including the Epigenetic Approaches in Cancer Epidemiology programme, which funds about 30 projects at a total of \$45 million. In 2011, the US National Institute of Environmental Health Sciences awarded about \$11 million in grants for epigenetics-related research. The institute has a strong interest in the environment's effect on epigenetics and how that influences diseases such as cancer, notes Edward Kang, a spokesman for the institute, which is based in Research Triangle Park, North Carolina.

Government investment has also fuelled the shift to large, genome-wide epigenomics studies. In October 2011, the European Commission launched its €39.9-million (US\$52.1-million) BLUEPRINT project, which brings together 41 institutes and companies to generate at least 100 reference epigenomes from healthy and leukaemic cells. Just over €2 million of that is still to be doled out, says project coordinator Henk Stunnenberg of Radboud University in Nijmegen, the Netherlands. The project's team hopes to recruit at least five more groups from academia and industry. The European Commission support of epigenetics research helped to woo Manel Esteller, an epigenetics researcher at the Bellvitge Biomedical Research Institute in Barcelona, Spain, back to his home country from the United States. Esteller now participates in the BLUEPRINT project and coordinates CURELUNG, another programme funded by the European Commission, which unites 11 institutions and companies and has analysed DNA methylation in nearly 1,000 human lung tumours thus far. "The European Commission offered the opportunity to apply for different grants that were able to complement local funding," he says. "This extra help has been critical in the success of my projects."

Many of the biggest investments in

epigenomics directly fund the larger sequencing centres rather than individual investigators, but smaller laboratories have capitalized on the steady stream of data and new technologies emerging from the programmes. In 2008, the US National Institutes of Health (NIH) launched a \$200-million, ten-year Roadmap Epigenomics Project to develop mapping centres and technologies that would allow researchers to survey epigenetic changes on a genome-wide scale. Although the project focuses on the epigenetics of normal, non-cancerous tissue, the technological advances and large data sets have helped cancer research as well. Many cancer genome sequencing projects, including the NCI's The Cancer Genome Atlas (TCGA) programme, include a partial focus on cataloguing epigenetic changes. Kenna Shaw, director of the TCGA programme office in Bethesda, says that the programme has funded around 200 jobs. The bulk of the funding for these large-scale programmes is already dedicated to the larger sequencing centres, but smaller teams are using the data from these projects to generate individual-investigator grant applications, Shaw adds.

These data have helped to persuade investors in industry that epigenetic abnormalities in cancer could provide a wealth of new drug targets. The finding that mutations in epigenetics-related genes may be driving some cancers offers the tantalizing possibility of taking a personalized approach to cancer treatment, a tack that is rapidly gaining



Personalized treatment for cancer is rapidly gaining ground in industry.

Robert Gould

ground in industry, says Robert Gould, chief executive of Epizyme, an epigenetics focused biotechnology firm based in Cambridge, Massachusetts. This evidence, plus the successful approval of a first generation of drugs intended to target epigenetic pathways, has convinced almost every major drug company to invest in cancer epigenetics, says Mukesh Verma, a programme officer at the NCI. For example, Novartis, a pharmaceutical firm with its headquarters in Basel, Switzerland, has more than 200 employees working in epigenetics, most of them in cancer, says En Li, head of China Novartis Institutes for Biomedical Research, based in Shanghai. Last year, GlaxoSmithKline in London, in addition to funding its own epigenetics team, paid \$20 million to partner with Epizyme in a deal

L. JIN/EPZYME

M. MCKEE/EPZYME

in which Epizyme could ultimately receive as much as \$630 million. “GSK’s group is partnering with us and is also competing with us on other programmes,” says Epizyme’s chief scientific officer, Robert Copeland. “It makes for an interesting dynamic.”

With so much excitement, competition in the field can be fierce. Data from large government projects can be a boon to smaller labs, says Clark, but individual investigators and those new to the field need to carve their own niche. “In the face of those big initiatives, smaller labs have the challenge of asking smaller and more unique questions as to the basic mechanisms underlying these epigenetic changes,” she says. Christopher Vakoc, an epigenetics researcher at Cold Spring Harbor Laboratory in New York, notes that the “tiny” lab he started in 2008 directly competed with several big pharmaceutical companies to discover a role for Brd4 — a ‘reader’ protein that binds to certain modified histones and modulates gene expression — in acute myeloid leukaemia (J. Zuber *et al. Nature* **478**, 524–528; 2011). After his team’s paper was published, Vakoc heard rumours that ten companies were racing to capitalize on the results.

There is also an intense demand for talent. In particular, epigenetics companies and individual labs need bioinformaticians as sequencing projects continue to dump terabytes of data into public databases (see *Nature* **482**, 263–265; 2012). Although this is an opportunity for job hunters with computational training, it creates challenges for those opening labs for the first time, says Jun Song, a computational biologist who opened his lab at the University of California, San Francisco, in 2009. Song has struggled to compete with bigger labs to recruit graduate students and postdoctoral researchers, who often prefer the proven track-record and extensive connections offered by a well-established principal investigator. “We battle to get a talented bioinformatician,” says Clark. “Everybody wants their own.”

Ultimately, Song looked outside biology to recruit three postdocs, two of whom he lured away from high-energy particle physics and the third from applied mathematics. Song himself was trained as a physicist, and says that epigenetics and epigenomics offer a range of challenging computational questions that can entice researchers from other fields. “It would be great to have someone already trained in both biology and computation,” he says. “But as biology becomes more quantitative as a field, I also believe that it’s important to bring in new computational scientists and train them in biology.”

The opportunity for cross-disciplinary training in epigenetics can be an advantage for bioinformaticians and molecular biologists alike, says Garcia. “It makes you a more well-rounded scientist,” he says. “And that’s what you need these days to compete in the job market.” ■

Heidi Ledford writes for *Nature* from Cambridge, Massachusetts.

COLUMN

A tough climb

Challenging your own ideas and opinions takes more than just a change of scenery, says **Andrew Peterman**.

It was 4:00 a.m., and I was sure I was getting close to the top. The wind had pelted my face with snow and ice for the past three hours. Every few steps, the train of people stopped. Below me, hundreds of specks of light from climbers’ lamps clung to the mountainside in a zigzag pattern. At each pause, I shut my eyes.

When I opened them again, I was looking down at the half-metre between my feet and the heels of my former college roommate. The short respite hardly counteracted the fact that each breath contained less than half of the oxygen I am used to back at home. I looked at my altimeter — I still had a couple of hours to go.

Last February, I decided to climb Mount Kilimanjaro in Tanzania, which stands 5,895 metres above sea level. I embarked on the 3-week trip to challenge myself to embrace a different culture. But I found that it takes more than a change of scenery to challenge one’s perceptions.

I wanted to broaden my landscape, test my own conventions and walk away feeling as if I had pushed myself physically and mentally. I wanted to create an unconventional forum for discussion, as different as possible from that of the engineering department at Stanford University, California. I invited my closest friends who had gone on to pursue different areas of study or practice from my own. In academia, we often interact with the same people, hear and speak the same language, and attend the same presentations. We surround ourselves with people just like ourselves. I assumed that an unfamiliar location and culture would challenge my ideas and opinions.

But researchers such as Miller McPherson, a sociologist at Duke University in Durham, North Carolina, have shown that similarity breeds connection — the homophily principle (M. McPherson *et al. Annu. Rev. Sociol.* **27**, 415–444; 2001). Individuals’ relationships tend towards homogeneity. In other words, we develop contacts with greater frequency among individuals who have sociodemographic and behavioural characteristics and attitudes similar to our own.

Despite the fact that my friends have pursued careers in other fields, they are still more like me than are other people. We are all males and are mostly white, Stanford alumni, from middle-upper-class families, in our late 20s who share similar political views. Perhaps



R. STEIER

Andrew Peterman and friends climb Kilimanjaro.

forming the group was, by my own subconscious design, a way to avoid the unfamiliar in a trying and scary environment, and perhaps the research is correct.

The experience has made me realize that homophily is also a tough mountain to overcome. I found that by stepping outside my comfort zone physically — braving the cold, harsh conditions of Kilimanjaro — I had clung to the familiar opinions of my close friends.

As much of the research in this area shows, homophily has serious implications for the development of new ideas. If you surround yourself with people who share your opinions, attitudes, beliefs and even experiences, how can you learn anything new? Who will challenge your ideas?

I aim to keep looking for that interdisciplinary environment. The first step is engaging with people with whom I do not always agree — embracing the conflict and uncomfortable nature of working with those with starkly different opinions. I believe that all scientists, especially those with interdisciplinary aspirations, should strive to break away from the familiar in search of the unfamiliar. Doing so may uncover a new approach to an old problem.

Creating these situations requires an active effort to push through the discomfort of difference. And, despite what the research suggests, it does not always have to be the case that ‘birds of a feather flock together’. ■

Andrew Peterman is a doctoral candidate in civil engineering at Stanford University in California.

DIVERSITY

PhD completion rates

In the hope of boosting degree completion rates, the US Council of Graduate Schools (CGS) in Washington DC is to examine attrition of minority students in science, technology, engineering and maths (STEM) programmes. The CGS will analyse data from 21 public and private universities for those entering programmes between 1992 and 2012. It aims to visit sites and interview students, faculty members and administrators to identify impediments to completion, and develop tools to remove them. Previous studies found that completion rates of minority students for STEM PhDs were significantly lower than those of non-minority students, notes Robert Sowell, vice-president for programmes and operations at CGS.

UNITED STATES

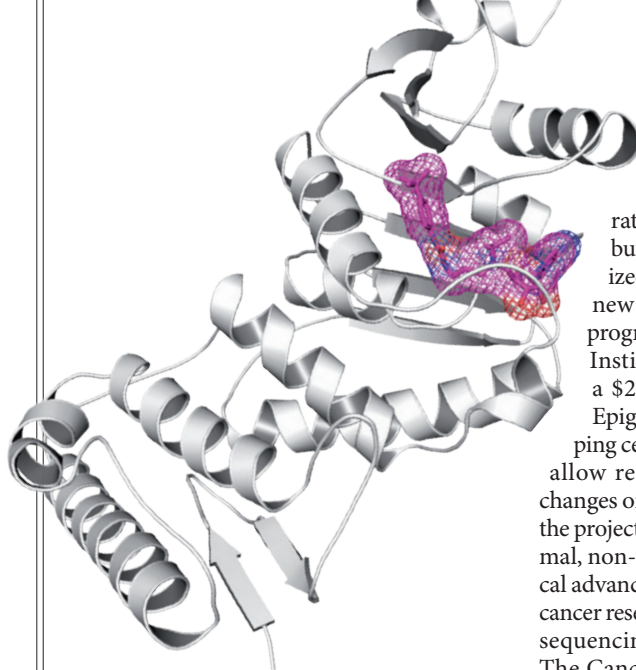
Unions banned

Michigan Governor Rick Snyder has banned graduate-student research assistants in public universities from unionizing following the efforts of 1,200 students to organize a union in April 2011. Snyder said in a statement that research assistants are students and giving them public-employee status and union representation would alter the student-teacher relationship. This is the latest action against US graduate-student unions. In 2004, New York University's union was disbanded under a state labour-board decision. Student representatives from Michigan State University in East Lansing and University of Michigan in Ann Arbor did not respond to interview requests.

PARTNERSHIPS

Postdoc opportunities

The California Institute for Quantitative Biosciences (QB3), part of the University of California, will hire up to 15 postdocs in a collaboration with drug firm Pfizer that expands a 2009 agreement to discover and develop technologies and drugs. Postdocs will be funded for two years in areas such as cardiovascular disease, immunology, neuroscience and oncology. They will learn to work with industry, says QB3 director Regis Kelly, who notes that this is a key activity given that many will go on to seek industry positions. Pfizer contributed US\$9.5 million to the original partnership and will provide at least the same level of funding again, says Ron Newbold, Pfizer's vice-president for strategic research partnerships.



An example of one of Epizyme's inhibitors interacting with an epigenetic enzyme.

► Cancer Institute (NCI) in Bethesda, Maryland, has several programmes dedicated to epigenetics, including the Epigenetic Approaches in Cancer Epidemiology programme, which funds about 30 projects at a total of \$45 million. In 2011, the US National Institute of Environmental Health Sciences awarded about \$11 million in grants for epigenetics-related research. The institute has a strong interest in the environment's effect on epigenetics and how that influences diseases such as cancer, notes Edward Kang, a spokesman for the institute, which is based in Research Triangle Park, North Carolina.

Government investment has also fuelled the shift to large, genome-wide epigenomics studies. In October 2011, the European Commission launched its €39.9-million (US\$52.1-million) BLUEPRINT project, which brings together 41 institutes and companies to generate at least 100 reference epigenomes from healthy and leukaemic cells. Just over €2 million of that is still to be doled out, says project coordinator Henk Stunnenberg of Radboud University in Nijmegen, the Netherlands. The project's team hopes to recruit at least five more groups from academia and industry. The European Commission support of epigenetics research helped to woo Manel Esteller, an epigenetics researcher at the Bellvitge Biomedical Research Institute in Barcelona, Spain, back to his home country from the United States. Esteller now participates in the BLUEPRINT project and coordinates CURELUNG, another programme funded by the European Commission, which unites 11 institutions and companies and has analysed DNA methylation in nearly 1,000 human lung tumours thus far. "The European Commission offered the opportunity to apply for different grants that were able to complement local funding," he says. "This extra help has been critical in the success of my projects."

Many of the biggest investments in

epigenomics directly fund the larger sequencing centres rather than individual investigators, but smaller laboratories have capitalized on the steady stream of data and new technologies emerging from the programmes. In 2008, the US National Institutes of Health (NIH) launched a \$200-million, ten-year Roadmap Epigenomics Project to develop mapping centres and technologies that would allow researchers to survey epigenetic changes on a genome-wide scale. Although the project focuses on the epigenetics of normal, non-cancerous tissue, the technological advances and large data sets have helped cancer research as well. Many cancer genome sequencing projects, including the NCI's The Cancer Genome Atlas (TCGA) programme, include a partial focus on cataloguing epigenetic changes. Kenna Shaw, director of the TCGA programme office in Bethesda, says that the programme has funded around 200 jobs. The bulk of the funding for these large-scale programmes is already dedicated to the larger sequencing centres, but smaller teams are using the data from these projects to generate individual-investigator grant applications, Shaw adds.

These data have helped to persuade investors in industry that epigenetic abnormalities in cancer could provide a wealth of new drug targets. The finding that mutations in epigenetics-related genes may be driving some cancers offers the tantalizing possibility of taking a personalized approach to cancer treatment, a tack that is rapidly gaining



Personalized treatment for cancer is rapidly gaining ground in industry.

Robert Gould

ground in industry, says Robert Gould, chief executive of Epizyme, an epigenetics focused biotechnology firm based in Cambridge, Massachusetts. This evidence, plus the successful approval of a first generation of drugs intended to target epigenetic pathways, has convinced almost every major drug company to invest in cancer epigenetics, says Mukesh Verma, a programme officer at the NCI. For example, Novartis, a pharmaceutical firm with its headquarters in Basel, Switzerland, has more than 200 employees working in epigenetics, most of them in cancer, says En Li, head of China Novartis Institutes for Biomedical Research, based in Shanghai. Last year, GlaxoSmithKline in London, in addition to funding its own epigenetics team, paid \$20 million to partner with Epizyme in a deal

L. JIN/EPZYME

M. MCKEE/EPZYME

in which Epizyme could ultimately receive as much as \$630 million. “GSK’s group is partnering with us and is also competing with us on other programmes,” says Epizyme’s chief scientific officer, Robert Copeland. “It makes for an interesting dynamic.”

With so much excitement, competition in the field can be fierce. Data from large government projects can be a boon to smaller labs, says Clark, but individual investigators and those new to the field need to carve their own niche. “In the face of those big initiatives, smaller labs have the challenge of asking smaller and more unique questions as to the basic mechanisms underlying these epigenetic changes,” she says. Christopher Vakoc, an epigenetics researcher at Cold Spring Harbor Laboratory in New York, notes that the “tiny” lab he started in 2008 directly competed with several big pharmaceutical companies to discover a role for Brd4 — a ‘reader’ protein that binds to certain modified histones and modulates gene expression — in acute myeloid leukaemia (J. Zuber *et al. Nature* **478**, 524–528; 2011). After his team’s paper was published, Vakoc heard rumours that ten companies were racing to capitalize on the results.

There is also an intense demand for talent. In particular, epigenetics companies and individual labs need bioinformaticians as sequencing projects continue to dump terabytes of data into public databases (see *Nature* **482**, 263–265; 2012). Although this is an opportunity for job hunters with computational training, it creates challenges for those opening labs for the first time, says Jun Song, a computational biologist who opened his lab at the University of California, San Francisco, in 2009. Song has struggled to compete with bigger labs to recruit graduate students and postdoctoral researchers, who often prefer the proven track-record and extensive connections offered by a well-established principal investigator. “We battle to get a talented bioinformatician,” says Clark. “Everybody wants their own.”

Ultimately, Song looked outside biology to recruit three postdocs, two of whom he lured away from high-energy particle physics and the third from applied mathematics. Song himself was trained as a physicist, and says that epigenetics and epigenomics offer a range of challenging computational questions that can entice researchers from other fields. “It would be great to have someone already trained in both biology and computation,” he says. “But as biology becomes more quantitative as a field, I also believe that it’s important to bring in new computational scientists and train them in biology.”

The opportunity for cross-disciplinary training in epigenetics can be an advantage for bioinformaticians and molecular biologists alike, says Garcia. “It makes you a more well-rounded scientist,” he says. “And that’s what you need these days to compete in the job market.” ■

Heidi Ledford writes for *Nature* from Cambridge, Massachusetts.

COLUMN

A tough climb

Challenging your own ideas and opinions takes more than just a change of scenery, says **Andrew Peterman**.

It was 4:00 a.m., and I was sure I was getting close to the top. The wind had pelted my face with snow and ice for the past three hours. Every few steps, the train of people stopped. Below me, hundreds of specks of light from climbers’ lamps clung to the mountainside in a zigzag pattern. At each pause, I shut my eyes.

When I opened them again, I was looking down at the half-metre between my feet and the heels of my former college roommate. The short respite hardly counteracted the fact that each breath contained less than half of the oxygen I am used to back at home. I looked at my altimeter — I still had a couple of hours to go.

Last February, I decided to climb Mount Kilimanjaro in Tanzania, which stands 5,895 metres above sea level. I embarked on the 3-week trip to challenge myself to embrace a different culture. But I found that it takes more than a change of scenery to challenge one’s perceptions.

I wanted to broaden my landscape, test my own conventions and walk away feeling as if I had pushed myself physically and mentally. I wanted to create an unconventional forum for discussion, as different as possible from that of the engineering department at Stanford University, California. I invited my closest friends who had gone on to pursue different areas of study or practice from my own. In academia, we often interact with the same people, hear and speak the same language, and attend the same presentations. We surround ourselves with people just like ourselves. I assumed that an unfamiliar location and culture would challenge my ideas and opinions.

But researchers such as Miller McPherson, a sociologist at Duke University in Durham, North Carolina, have shown that similarity breeds connection — the homophily principle (M. McPherson *et al. Annu. Rev. Sociol.* **27**, 415–444; 2001). Individuals’ relationships tend towards homogeneity. In other words, we develop contacts with greater frequency among individuals who have sociodemographic and behavioural characteristics and attitudes similar to our own.

Despite the fact that my friends have pursued careers in other fields, they are still more like me than are other people. We are all males and are mostly white, Stanford alumni, from middle-upper-class families, in our late 20s who share similar political views. Perhaps



Andrew Peterman and friends climb Kilimanjaro.

forming the group was, by my own subconscious design, a way to avoid the unfamiliar in a trying and scary environment, and perhaps the research is correct.

The experience has made me realize that homophily is also a tough mountain to overcome. I found that by stepping outside my comfort zone physically — braving the cold, harsh conditions of Kilimanjaro — I had clung to the familiar opinions of my close friends.

As much of the research in this area shows, homophily has serious implications for the development of new ideas. If you surround yourself with people who share your opinions, attitudes, beliefs and even experiences, how can you learn anything new? Who will challenge your ideas?

I aim to keep looking for that interdisciplinary environment. The first step is engaging with people with whom I do not always agree — embracing the conflict and uncomfortable nature of working with those with starkly different opinions. I believe that all scientists, especially those with interdisciplinary aspirations, should strive to break away from the familiar in search of the unfamiliar. Doing so may uncover a new approach to an old problem.

Creating these situations requires an active effort to push through the discomfort of difference. And, despite what the research suggests, it does not always have to be the case that ‘birds of a feather flock together’. ■

Andrew Peterman is a doctoral candidate in civil engineering at Stanford University in California.

R. STEIER

INVISIBLE

The path to immortality.

BY JOÃO RAMALHO-SANTOS

He slid out of bed as the door closed behind the nurse who regularly came by to check if he was still breathing. Avoidance was always best; unlike academia, this was a place where quick wits were greeted, not by admiration, but with increased doses of meds. Keeping them controlled was the only goal. Nurses weren't impressed by who their charges had been; they dealt with ex-politicians, ex-actors, ex-chief executives, ex-everything, focus on the 'ex'. The trick was to be invisible, to walk the fine line between polite privacy and antisocial sullenness. Rather than musing on 'how it had come to this', he took it for what it was: a new challenge.

Today, however, the wait had been excruciating, a package beckoning just outside the door. The nurse never brought the mail in, not part of the job description. But it was there; he knew it, next-day shipping never failed. Fifteenth edition. Two shelves on the bookcase held the fourteen previous ones, a steady increase in bulk following the chronology. In fact, these were the only books he had bothered to bring. He opened the door, trying to will away telltale creaks in hinges and joints, avoid any possible attention. But a small envelope was all that awaited.

A sudden surge of adrenaline-flavoured fear gushed through him. The publishing company had gone all-digital. Inside the envelope would be a DVD, a USB pen, a code to access some website far away. No longer the heaviness of textbooks, the rustle of knowledge to be thumbed through, the smell of fresh ink; just jumps, links and animations, information beaten into easy morsels. Yet another challenge, he mused, firing up the laptop, searching for glasses, battling arthritis for the envelope's contents.

The chapter was not where he expected; the new authors had wanted to shift things around, leave their mark. Wouldn't work: by now the book was known by a sole last name, and that original author had been dead since the tenth edition, his name transitioning from scholar to brand. But even creative authors couldn't escape the obvious organizations in science, he thought, finding what he was looking for.

➔ **NATURE.COM**
Follow Futures on
Facebook at:
go.nature.com/mtodm

One introductory line.

"It has long

been well established that..." No references were given. The chapter then proceeded to describe what had recently happened in the field. Why, the new authors must have thought, reference the obvious at the beginning? They had merely added what seemed like a million links at the end, for those with a taste for the historical. He grinned, gazed at the bookcase.

The first four editions he forgave, only the drive for completeness justified their purchase. He was in high school when the first two came out, in college for the others. The fifth he had learned to understand. When it was published he had only presented at a meeting, and at the time hadn't even been fully aware of what the data meant. It was the sixth and seventh he had real issues with. By then his PhD thesis had been completed, the data published, their implications clear. Yet it remained ignored, just a few odd details that didn't quite fit accepted dogma, certainly not enough to warrant the rewriting of textbooks, as one helpful professor candidly explained. So he formed his own lab to

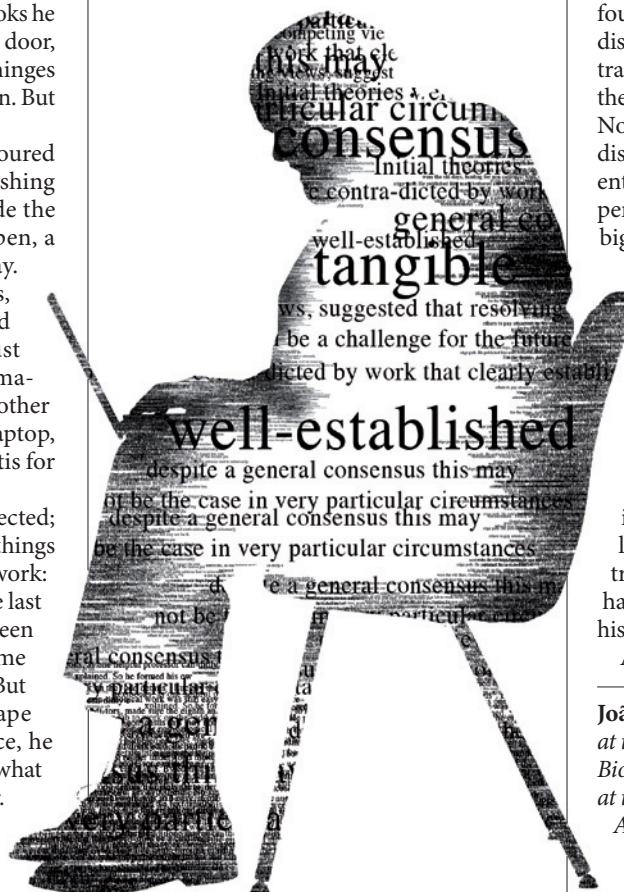
work on the 'odd details'. Luckily these were the old days, funding for non-canonical work was still easy, if off the beaten prestige path. He published like mad, bothered editors, made sure the eighth and ninth editions had to reluctantly state: "Despite a general consensus this may not be the case in very particular circumstances." Finally he was referenced, the work tangible; even though any casual reader understood the textbook was being, at best, charitable. By edition number ten his relentless campaign had got others to pay attention, to try out his hypotheses. No longer the ramblings of a lone maverick, the text finally admitted that there were competing views, suggested that resolving this issue would be a challenge for the future.

And the future came through in editions eleven to thirteen, his work gradually becoming the "general consensus", the previous fading into afterthought. The thirteenth edition was particularly satisfying because he had since retired, the ideas no longer dependent on his own stubbornness, but on the best truth available.

Five years ago, when he first read the fourteenth, he had to admit to a twinge of disappointment. "Initial theories were contradicted by work that clearly established..." the chapter said, still referencing his papers. Nothing else. It was as if the fiery battles discussed in previous editions, and that his entire career was based upon, hadn't happened at all. But slowly he understood the bigger picture, realized what the next edition, what all future versions, would have to say.

And fifteen did. The controversy was dead, to resurface in other chapters on the history of the field, but not useful in day-to-day practice, realm of the "well-established". Later he would check if Wikipedia and Google Scholar agreed, but the grin was already turning into his first real smile in years. Regardless of all the awards and accolades, the true pinnacle of the academic profession had now been reached. Peers considered his work good enough to be truly immortal. And here too he was finally invisible. ■

João Ramalho-Santos has been sighted at the Center for Neuroscience and Cell Biology and the Department of Life Sciences at the University of Coimbra, Portugal. And several other places. He likes them all equally, but when he is in one, he often wishes he were in another.



JACEY

Air density 2.7 billion years ago limited to less than twice modern levels by fossil raindrop imprints

Sanjoy M. Som^{1,2†}, David C. Catling¹, Jelte P. Harnmeijer^{1,3}, Peter M. Polivka^{1,4} & Roger Buick¹

According to the ‘Faint Young Sun’ paradox, during the late Archaean eon a Sun approximately 20% dimmer warmed the early Earth such that it had liquid water and a clement climate¹. Explanations for this phenomenon have invoked a denser atmosphere that provided warmth by nitrogen pressure broadening¹ or enhanced greenhouse gas concentrations². Such solutions are allowed by geochemical studies and numerical investigations that place approximate concentration limits on Archaean atmospheric gases, including methane, carbon dioxide and oxygen^{2–7}. But no field data constraining ground-level air density and barometric pressure have been reported, leaving the plausibility of these various hypotheses in doubt. Here we show that raindrop imprints in tuffs of the Ventersdorp Supergroup, South Africa, constrain surface air density 2.7 billion years ago to less than twice modern levels. We interpret the raindrop fossils using experiments in which water droplets of known size fall at terminal velocity into fresh and weathered volcanic ash, thus defining a relationship between imprint size and raindrop impact momentum. Fragmentation following raindrop flattening limits raindrop size to a maximum value independent of air density, whereas raindrop terminal velocity varies as the inverse of the square root of air density. If the Archaean raindrops reached the modern maximum measured size, air density must have been less than 2.3 kg m^{−3}, compared to today’s 1.2 kg m^{−3}, but because such drops rarely occur, air density was more probably below 1.3 kg m^{−3}. The upper estimate for air density renders the pressure broadening explanation¹ possible, but it is improbable under the likely lower estimates. Our results also disallow the extreme CO₂ levels required for hot Archaean climates⁸.

Numerical investigations of Archaean atmospheric composition^{2–6} typically assume a modern, total atmospheric pressure of about one atmosphere (1 atm), but there are good reasons why barometric pressure may have been different. First, the partial pressure of oxygen p_{O_2} was negligible before the Great Oxidation Event at around 2.35 billion years ago⁹. There are several independent lines of evidence for this¹⁰, the strongest being widespread and large mass-independent fractionations of sulphur isotopes in Archaean sediments that arise only in an atmosphere with less than about one part oxygen per million by volume (p.p.m.v.)¹¹. Second, before the advent of an aerobic nitrogen cycle coincident with rising oxygen levels¹², the flux of nitrogen back to the atmosphere via the now-dominant nitrification–denitrification pathway would have been different from now. So a lack of oxygen before the Great Oxidation Event should have affected the partial pressure of nitrogen p_{N_2} , the major gas contributing to total atmospheric pressure. Moreover, it has been calculated that a p_{N_2} of 2.37 atm at 2.5 billion years ago could solve the ‘Faint Young Sun’ paradox by pressure-broadening infrared absorption of greenhouse gases¹. Other studies postulate a hot (~70 °C) Archaean ocean based on oxygen isotopes in cherts¹³, requiring a partial pressure of carbon dioxide p_{CO_2} of about 2–6 bar (ref. 8), which would contradict the p_{CO_2} levels of only 10–50 present atmospheric levels (PAL) constrained

from 2.69-billion-year-old palaeosols⁷. Such ambiguities concerning the composition of the ancient atmosphere could be resolved, or improved upon, by knowledge of total atmospheric pressure. Here, we use raindrop imprints to constrain total ground-level atmospheric density (and thus total surface pressure) 2.7 billion years before present. The idea of using raindrop imprints as a proxy for air density was suggested by Lyell¹⁴ in 1851 but has hitherto been unexplored.

On the ancient Earth, maximum raindrop diameters should have been essentially identical to today’s, because the maximum size beyond which raindrops disintegrate at terminal velocity is independent of air density. Falling raindrops flatten^{15,16} and fragment when the total aerodynamic forces exceed the combination of surface tension and hydrostatic forces¹⁷. Fragmentation begins when the raindrop bottom becomes flat at a force balance given by¹⁷:

$$V_{\text{term}}^2 d = \frac{8\gamma}{\rho_{\text{air}} n C_d} \quad (1)$$

where V_{term} is terminal velocity, d is the diameter of a sphere equivalent to the drop volume, γ is surface tension, ρ_{air} is air density, C_d is the drag coefficient, and n is a factor relating the radius of the upper curvature of the drop to its spherical equivalent radius. Theory relates terminal velocity to raindrop size^{18,19} and predicts 9.3 m s^{−1} for a raindrop of 6.8 mm in diameter, the largest measured raindrop at ground level²⁰. Typical values under standard surface atmospheric conditions ($\gamma = 7 \times 10^{-2}$ N m^{−1} and $\rho_{\text{air}} = 1.2$ kg m^{−3}) yield a constant value of 0.80 for $n C_d$ in equation (1), comparing favourably with $n C_d = 0.85$ from independent studies¹⁷ and consistent with observations that the product $V_{\text{term}}^2 d$ is constant^{21,22}.

A further relationship derived from empirical correlations exists between air density and maximal terminal velocity^{18,23,24}:

$$V_{\text{term,max}} = 2 \left(\frac{\rho_{\text{water}} g \gamma}{\rho_{\text{air}}^2} \right)^{0.25} \approx \frac{\text{a known constant}}{\rho_{\text{air}}^{0.5}} \quad (2)$$

where g is gravitational acceleration and ρ_{water} is the density of water. Equation (2) also gives a maximum terminal velocity of 9.3 m s^{−1}, which corresponds to the largest raindrops of 6.8 mm in diameter. Substitution of $V_{\text{term}} \propto \rho_{\text{air}}^{-1/2}$ from equation (2) into $d \propto (V_{\text{term}}^2 \rho_{\text{air}})^{-1}$ from equation (1) cancels out ρ_{air} , showing that maximum raindrop size is independent of air density. Drop equivalent diameter d is thus simply a function of surface tension γ . The slight increase of surface tension with temperature causes only a trivial terminal velocity change of 0.05 m s^{−1} over 15–30 °C, meaning that somewhat different Archaean temperatures would not affect our conclusions. Consequently, an upper bound on air density can be derived from the largest raindrop imprints, formed by the transfer of momentum from the largest impacting raindrops to the substrate.

The Archaean imprints studied here (Fig. 1 and Supplementary Information) are from the Omdraaiivlei farm near Prieska, South Africa, in the Kameeldoorns Formation of the Platberg Group, the middle unit of the 2.7-billion-year-old Ventersdorp Supergroup²⁵

¹Department of Earth and Space Sciences and Astrobiology Program, University of Washington, Seattle, Washington 98195-1310, USA. ²Blue Marble Space Institute of Science, Seattle, Washington 98145, USA. ³Sustainable Community Energy Network, Edinburgh Centre for Low Carbon Innovation, Edinburgh EH8 9AA, UK. ⁴Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington 98195, USA. [†]Present address: Exobiology Branch, NASA Ames Research Center, Moffett Field, California 94035, USA.

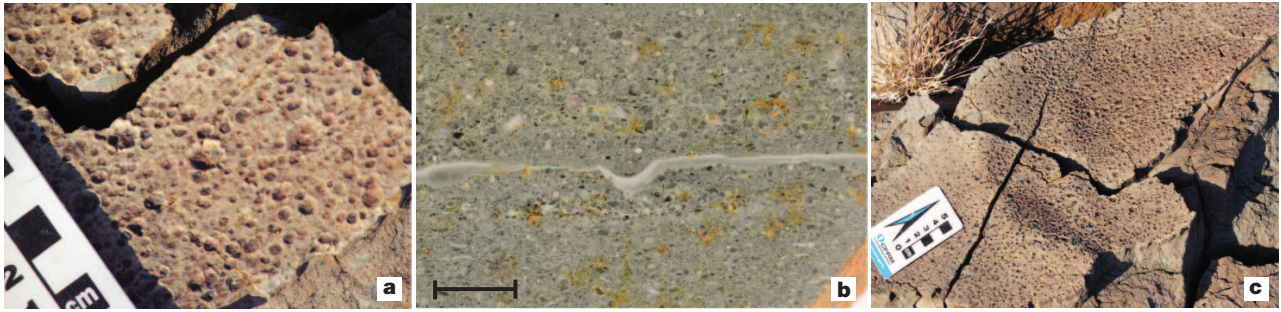


Figure 1 | The 2.7-billion-year-old Ventersdorp Supergroup raindrop imprints lithified in tuff at Omdraaivlei, South Africa. a, Detail of slightly elliptical outlines of raindrop imprints. **b,** Cross-section photograph of imprints penetrating 1–2 mm into coarse accretionary lapilli tuff, and draped

(Supplementary Fig. 2). They penetrate into very poorly sorted fine tuff of intermediate igneous composition. A layer of pale, very fine volcanic ash 0.5–0.8 mm thick drapes the imprints (Fig. 1b and Supplementary Information), which reduces the diameter of the exposed imprints relative to their original diameter by about twice the drape thickness. The rimmed craters are well preserved, slightly elliptical in shape and occasionally overlap, suggesting that the rain event that formed them was of short duration and of light to moderate intensity, because high-intensity rainfall leads to distorted imprints and long-duration showers cause substantial overlap²⁶. They were probably formed in an inland semi-arid setting near sea level (Supplementary Information).

The dominant contributor to imprint size is the change in raindrop momentum during impact²⁷. We define a dimensionless momentum J as:

$$J = \frac{V_{\text{term}} m_d}{\eta A_d} \quad (3)$$

where m_d is the mass of the raindrop, η is the dynamic viscosity (independent of ρ_{air}), and A_d is the cross-sectional area of the falling drop. Obtaining atmospheric density from lithified raindrop imprints requires: (1) measuring raindrop imprint areas; (2) determining experimentally how J varies with imprint area by varying d and thus m_d in equation (3) (Fig. 2); and (3) relating atmospheric density to J (Fig. 3). Archaean imprints were measured in the field, and subsequently re-measured by high-resolution three-dimensional laser scanning of latex peels taken *in situ*. The relationship between drop impact momentum and corresponding imprint area was obtained from experiments in which we released water drops of known mass from an indoor height sufficient to guarantee that they reached terminal velocity onto ash substrates analogous to the Archaean tuff. One of the experimental ashes was fresh from the 2010 Eyjafjallajökull

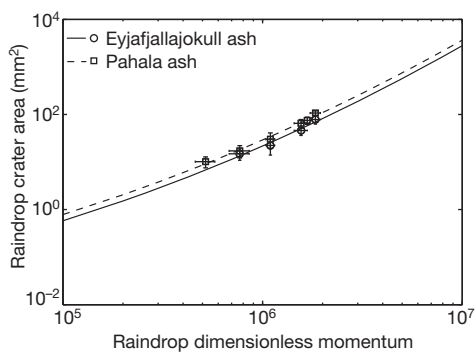


Figure 2 | Experimental relationship between raindrop area and dimensionless momentum. The horizontal error bars are the uncertainty of raindrop mass propagated to dimensionless momentum; the vertical error bars express the corresponding standard deviation of crater dimensions.

with a thin veneer (about 0.5 mm) of light-toned, fine volcanic ash. Scale bar, 10 mm. **c,** Mildly increased imprint density on the windward (north-facing) faces of underlying symmetrical wave-ripples. (Photo credits: W. Altermann for **a** and **c**, and T. Tobin for **b**).

eruption in Iceland, and the other was weathered late Pleistocene Pahala ash from Hawaii²⁸. Both were from mafic to intermediate hydro-volcanic eruptions with similar grain-size distributions to the ash hosting the Archaean raindrop imprints. The relationship between air density and dimensionless momentum was obtained by extrapolating from previous work relating air density to terminal velocity¹⁸ (Supplementary Information).

The actual raindrop diameter that formed the largest imprints found at Omdraaivlei is unknown. Smaller imprint area reflects lower raindrop velocity and thus higher air density. Figure 4 illustrates the expected atmospheric density when the raindrop size that caused the largest Omdraaivlei imprint is varied. To calculate an atmospheric density upper bound, we use the lower bound on the largest raindrop imprint area, because smaller imprint areas reflect lower raindrop terminal velocities and hence higher air density. The largest imprint area is bounded by $A_{\text{latex}} - 1\sigma$ and $A_{\text{in situ}} + 1\sigma$, where $A_{\text{in situ}}$ is the mean maximum imprint area measured in the field, and A_{latex} is the mean maximum imprint area measured from the latex peels (Supplementary Information). Finding the corresponding air density for these end-member dimensions for a fixed raindrop size defines the error in air density. The air density upper bound is calculated as $A_{\text{latex}} - 1\sigma$. Using a drop diameter of 6.8 mm—the size of the largest raindrop ever measured at ground level²⁰ and also the theoretical maximum size—we obtain an absolute upper limit of less than 2.3 kg m^{-3} . However, because rainfall events producing these maximal drops are extremely rare, very intense, and highly erosive²⁶, it is more likely that the maximum raindrop size responsible for creating the Archaean imprints had an equivalent diameter of 3.8–5.3 mm, depending upon the choice of parameterization of the raindrop size distribution and assuming that the probability distribution functions for rainfall rates in inland semi-arid settings were similar on the Archaean Earth and the modern Earth. These dimensions correspond to a more likely upper limit for atmospheric density of $0.6\text{--}1.3 \text{ kg m}^{-3}$.

Estimates of atmospheric pressure from the ideal gas law $P = \rho_{\text{air}} RT$ require assumptions about air temperature T and, through the specific gas constant R , atmospheric composition. Regarding temperature 2.7 ± 0.1 billion years ago, no evidence of glaciation is present in the rock record. This may reflect lack of preservation, but if it is real, an absence of glaciation requires average temperatures to have been 20°C or higher, according to data from non-glacial times in the Phanerozoic eon²⁹. This is also consistent with Archaean temperatures of less than 40°C based on oxygen isotope systematics³⁰. Taking a nominal temperature of about 20°C , we calculated an upper limit on atmospheric pressure by choosing a composition that maximizes R . A $100\% \text{ N}_2$ atmosphere ($R = 297 \text{ J kg}^{-1} \text{ K}^{-1}$, versus $R = 253 \text{ J kg}^{-1} \text{ K}^{-1}$ for a $70\% \text{ N}_2 + 30\% \text{ CO}_2$ atmosphere) constrains atmospheric pressure 2.7 billion years ago to below 0.52–1.1 atm if we take ρ_{air} as less than $0.6\text{--}1.3 \text{ kg m}^{-3}$ or an absolute upper limit of less than 2.1 atm if we take $\rho_{\text{air}} = 2.3 \text{ kg m}^{-3}$.

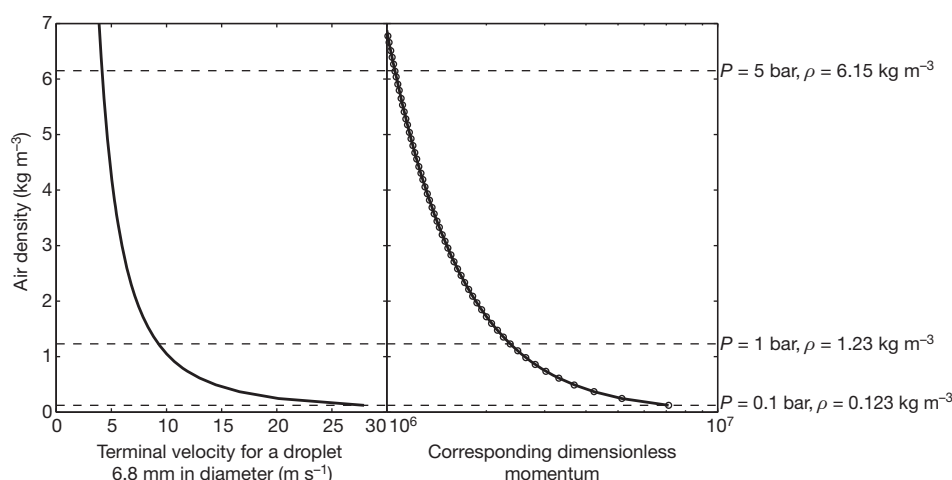


Figure 3 | Theoretical predictions of the variation of air density with terminal velocity and dimensionless momentum at the surface. a, The relationship between air density and terminal velocity, following the method of

ref. 18. **b,** The relationship between air density and dimensionless momentum (see Supplementary Information).

Our result extends quantifiable atmospheric pressure determinations beyond the modern era to the early Earth. It places constraints on some Archaean climate models^{1,13}, but does not invalidate other proposed late-Archaean atmospheres^{2–6}. For models invoking high p_{N_2} as a means of pressure-broadening greenhouse gas absorption¹, only the lowest estimate of 1.58 atm is close to our findings, suggesting a nitrogen content in the late-Archaean atmosphere of no more than twice present levels. Our result rules out very high Archaean ocean temperatures of 70 °C–85 °C (refs 13 and 31), because these would necessitate about 2–6 bar of carbon dioxide⁸ plus 0.3–0.6 bar of water vapour, increasing barometric pressure far beyond the upper limit found here. Thus, neither strong pressure-broadening nor extreme p_{CO_2} are satisfactory mechanisms for warming the early Earth illuminated by a 'Faint Young Sun'.

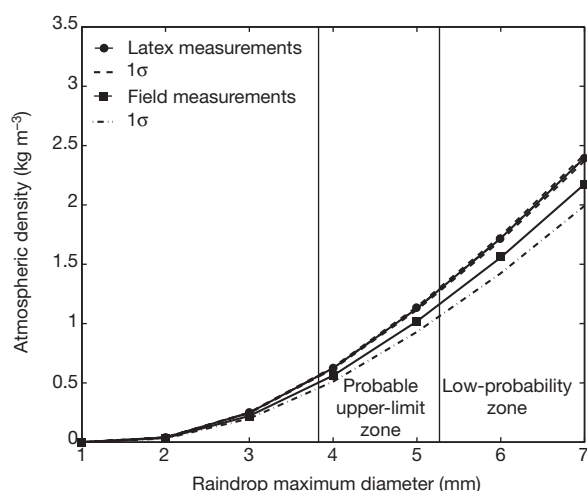


Figure 4 | Atmospheric density given the maximum raindrop diameter that created the largest imprints at Omdraaivlei, South Africa. Squares represent the air density calculated from the maximum imprint areas measured *in situ* (ash drape removed). Circles represent the air density calculated from the maximum imprint areas from latex measurement of imprint areas (ash drape removed). Dashed lines represent 1σ error. With the assumption that the rainfall rate probability distribution function responsible for the Archaean imprints is analogous to a modern, semi-arid, rainfall rate probability distribution function, there is a 78–99% probability that the maximum raindrop diameter was less than 3.8–5.3 mm (Supplementary Fig. 5, Supplementary Table 1, and Supplementary Information).

METHODS SUMMARY

We measured the Ventersdorp raindrop imprints directly in the field and by casting them using low-viscosity latex for later laboratory study. The resulting latex peels captured the dimensions of 955 individual raindrop imprints. The topography of the peels was measured using high-resolution three-dimensional laser scanning. The corresponding point-clouds (available at http://gis.ess.washington.edu/papers/Sanjay_Som_raindrops/) were interpolated using an inverse data-weighting scheme to create a digital elevation model. The digital elevation models were imported into a Geographical Information System and the dimensions of the imprints extracted. The dimensions were optimally binned, with the largest bin corresponding to the measurement of the largest imprints, and the dimension of each bin reflecting error in measurement.

To find the relationship between raindrop imprint dimension and dimensionless momentum, we released droplets of different (known) mass from a height of 27 m indoors onto analogous ash taken from Iceland and Hawaii. This height is more than double that required for all drops to reach terminal velocity. We could calculate the dimensionless momentum of all impacting raindrops because terminal velocity is predictable. The resulting imprinted ash substrates were lithified using hair spray and low-viscosity liquid urethane plastic. The dimensions were measured using the same laser scanner as that used for the latex peels. Each imprinted tray captured a dozen imprints originating from raindrops of the same mass, from which a mean and standard deviation were obtained.

We followed published methods¹⁹ to predict theoretically from first principles how raindrop terminal velocity changes with air density, and thus how dimensionless momentum changes with air density. Given the measurement of the largest Ventersdorp imprint, we obtained the corresponding dimensionless momentum of the impacting drop using our experimental relationship. By assuming the dimension of the raindrop responsible for the largest imprint (bounded by the maximum diameter of 6.8 mm), we quantified atmospheric density (Supplementary Information).

Received 8 July 2011; accepted 23 January 2012.

Published online 28 March 2012.

- Goldblatt, C. *et al.* Nitrogen-enhanced greenhouse warming on early Earth. *Nature Geosci.* **2**, 891–896 (2009).
- Kasting, J. Theoretical constraints on oxygen and carbon dioxide concentrations in the Precambrian atmosphere. *Precamb. Res.* **34**, 205–229 (1987).
- Ohmoto, H., Watanabe, Y. & Kumazawa, K. Evidence from massive siderite beds for a CO₂-rich atmosphere before 1.8 billion years ago. *Nature* **429**, 395–399 (2004).
- Claire, M., Catling, D. & Zahnle, K. Biogeochemical modelling of the rise in atmospheric oxygen. *Geobiology* **4**, 239–269 (2006).
- Domagal-Goldman, S., Kasting, J., Johnston, D. & Farquhar, J. Organic haze, glaciations and multiple sulfur isotopes in the Mid-Archaean Era. *Earth Planet. Sci. Lett.* **269**, 29–40 (2008).
- Haqq-Misra, J., Domagal-Goldman, S., Kasting, P. & Kasting, J. A revised, hazy methane greenhouse for the Archaean Earth. *Astrobiology* **8**, 1127–1137 (2008).
- Driese, S. G. *et al.* Neoproterozoic paleoweathering of tonalite and metabasalt: implications for reconstructions of 2.69 Ga early terrestrial ecosystems and paleoatmospheric chemistry. *Precamb. Res.* **189**, 1–17 (2011).

8. Kasting, J. & Howard, M. Atmospheric composition and climate on the early Earth. *Phil. Trans. R. Soc. B* **361**, 1733–1742 (2006).
9. Bekker, A. *et al.* Dating the rise of atmospheric oxygen. *Nature* **427**, 117–120 (2004).
10. Holland, H. D. The oxygenation of the atmosphere and oceans. *Phil. Trans. R. Soc. B* **361**, 903–915 (2006).
11. Farquhar, J., Bao, H. & Thiemens, M. Atmospheric influence of Earth's earliest sulfur cycle. *Science* **289**, 756–758 (2000).
12. Garvin, J. *et al.* Isotopic evidence for an aerobic nitrogen cycle in the latest Archean. *Science* **323**, 1045–1048 (2009).
13. Knauth, L. & Lowe, D. High Archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga Swaziland Supergroup, South Africa. *Geol. Soc. Am. Bull.* **115**, 566–580 (2003).
14. Lyell, C. On fossil rain-marks of the Recent, Triassic, and Carboniferous periods. *Q. J. R. Geol. Soc.* **7**, 238–247 (1851).
15. Spilhaus, A. Raindrop size, shape and falling speed. *J. Atmos. Sci.* **5**, 108–110 (1948).
16. Magono, C. On the shape of water drops falling in stagnant air. *J. Meteorol.* **11**, 77–79 (1954).
17. Matthews, J. & Mason, B. Electrification produced by the rupture of large water drops in an electric field. *Q. J. R. Meteorol. Soc.* **90**, 275–286 (1964).
18. Lorenz, R. The life, death and afterlife of a raindrop on Titan. *Planet. Space Sci.* **41**, 647–655 (1993).
19. Pruppacher, H. R. & Klett, J. D. *Microphysics of Clouds and Precipitation* (Kluwer Academic, 1997).
20. Willis, P. & Tattelman, P. Drop-size distribution associated with intense rainfall. *J. Appl. Meteorol.* **28**, 3–15 (1989).
21. Dodd, K. On the disintegration of water drops in an air stream. *J. Fluid Mech.* **9**, 175–182 (1960).
22. Beard, K. V. Velocity and shape of cloud and precipitation drops aloft. *J. Atmos. Sci.* **33**, 851–864 (1976).
23. Clift, R., Grace, J. R. & Weber, M. E. *Bubbles, Drops and Particles* Ch. 7 (Academic, 1978).
24. Foote, G. B. & du Toit, P. S. Terminal velocity of raindrops aloft. *J. Appl. Meteorol.* **8**, 249–253 (1969).
25. van der Westhuizen, W., Grobler, N., Looek, J. & Tordiffe, E. Raindrop imprints in the Late Archaean–Early Proterozoic Ventersdorp Supergroup, South Africa. *Sedim. Geol.* **61**, 303–309 (1989).
26. Reineck, H. & Singh, I. S. *Depositional Sedimentary Environments* 61 (Springer, 1980).
27. Huang, C., Bradford, J. M. & Cushman, J. H. A numerical study of raindrop impact phenomena: the elastic deformation case. *Soil Sci. Soc. Am. J.* **47**, 855–861 (1983).
28. Easton, R. Stratigraphy of Kilauea Volcano. In *Volcanism in Hawaii* (eds Decker, R., Wright, T. & Stauffer, P.) 243–260 (US Government Printing Office, US Geol. Surv. Prof. Pap. 1350, 1987).
29. Zachos, J., Dickens, G. & Zeebe, R. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283 (2008).
30. Hren, M., Tice, M. & Chamberlain, C. Oxygen and hydrogen isotope evidence for a temperate climate 3.42 billion years ago. *Nature* **462**, 205–208 (2009).
31. Knauth, L. Temperature and salinity history of the Precambrian ocean: implications for the course of microbial evolution. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **219**, 53–69 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by NASA Exobiology/Astrobiology grant NNX08AP56G. We thank W. Van der Westhuizen of the University of the Free State in South Africa, and E. and D. Jackson of Omdraaivlei for their assistance when sampling in the field. We also thank E. Stüeken, A. Chen and K. Huntington at the University of Washington for laboratory assistance, and the staff at Metron Corporation for data acquisition. XRF measurements were performed by the Washington State University Geoanalytical Laboratory. Funding and field logistics for the Iceland fieldwork was supported by the Coordination Action for Research Activities on life in Extreme Environments (CAREX), a project supported by the European Commission Seventh Framework Programme. Funding and field logistics for the Hawaiian fieldwork was supported by the University of Washington Department of Earth and Space Sciences, and its Geoclub. D.C.C. was also supported by NASA Exobiology/Astrobiology grant NNX10AQ90G.

Author Contributions D.C.C. conceived the research project and established maximum-raindrop-size terminal velocity dependence on air density, R.B. led the field work in South Africa, collected the latex peels and analysed the grain sizes of the Ventersdorp tuff, J.P.H. found additional geographic and stratigraphic occurrences of raindrop imprints while performing field work and helping collect latex peels, P.M.P. collected experimental data and measured the Eyjafjallajökull and Pahala ash grain sizes, and S.M.S. developed the method of dimensionless momentum, collected the volcanic ash from Hawaii and Iceland, analysed the data, and led the experimental work. S.M.S., R.B. and D.C.C. discussed results and prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.M.S. (sanjay.m.som@nasa.gov).

IFITM3 restricts the morbidity and mortality associated with influenza

Aaron R. Everitt¹, Simon Clare¹, Thomas Pertel², Sinu P. John², Rachael S. Wash¹, Sarah E. Smith¹, Christopher R. Chin², Eric M. Feeley², Jennifer S. Sims², David J. Adams¹, Helen M. Wise³, Leanne Kane¹, David Goulding¹, Paul Digard³, Verner Anttila¹, J. Kenneth Baillie^{4,5}, Tim S. Walsh⁵, David A. Hume⁴, Aarno Palotie¹, Yali Xue¹, Vincenza Colonna^{1,6}, Chris Tyler-Smith¹, Jake Dunning⁷, Stephen B. Gordon⁸, The GenISIS Investigators*, The MOSAIC Investigators*, Rosalind L. Smyth⁹, Peter J. Openshaw⁷, Gordon Dougan¹, Abraham L. Brass^{2,10} & Paul Kellam^{1,11}

The 2009 H1N1 influenza pandemic showed the speed with which a novel respiratory virus can spread and the ability of a generally mild infection to induce severe morbidity and mortality in a subset of the population. Recent *in vitro* studies show that the interferon-inducible transmembrane (IFITM) protein family members potently restrict the replication of multiple pathogenic viruses^{1–7}. Both the magnitude and breadth of the IFITM proteins' *in vitro* effects suggest that they are critical for intrinsic resistance to such viruses, including influenza viruses. Using a knockout mouse model⁸, we now test this hypothesis directly and find that IFITM3 is essential for defending the host against influenza A virus *in vivo*. Mice lacking *Ifitm3* display fulminant viral pneumonia when challenged with a normally low-pathogenicity influenza virus, mirroring the destruction inflicted by the highly pathogenic 1918 'Spanish' influenza^{9,10}. Similar increased viral replication is seen *in vitro*, with protection rescued by the re-introduction of *Ifitm3*. To test the role of IFITM3 in human influenza virus infection, we assessed the *IFITM3* alleles of individuals hospitalized with seasonal or pandemic influenza H1N1/09 viruses. We find that a statistically significant number of hospitalized subjects show enrichment for a minor *IFITM3* allele (SNP rs12252-C) that alters a splice acceptor site, and functional assays show the minor CC genotype *IFITM3* has reduced influenza virus restriction *in vitro*. Together these data reveal that the action of a single intrinsic immune effector, IFITM3, profoundly alters the course of influenza virus infection in mouse and humans.

IFITM3 was identified in a functional genomic screen as mediating resistance to influenza A virus, dengue virus and West Nile virus infection *in vitro*¹. However, the role of the IFITM proteins in antiviral immunity *in vivo* is unknown. Therefore, we infected mice that are homozygous for a disruptive insertion in exon 1 of the *Ifitm3* gene that abolishes its expression⁸ (*Ifitm3*^{−/−}) with a low-pathogenicity murine-adapted H3N2 influenza A virus (A/X-31). Low-pathogenicity strains of influenza do not normally cause extensive viral replication throughout the lungs, or cause the cytokine dysregulation and death typically seen after infection with highly pathogenic viral strains⁹, at the doses used (Fig. 1a). However, low-pathogenicity-infected *Ifitm3*^{−/−} mice became moribund, losing >25% of their original body weight and showing severe signs of clinical illness (rapid breathing, piloerection) 6 days after infection. In comparison, wild-type littermates shed <20% of their original body weight, before fully recovering (Fig. 1a, b). There was little difference in virus replication in the lungs during the first 48 h

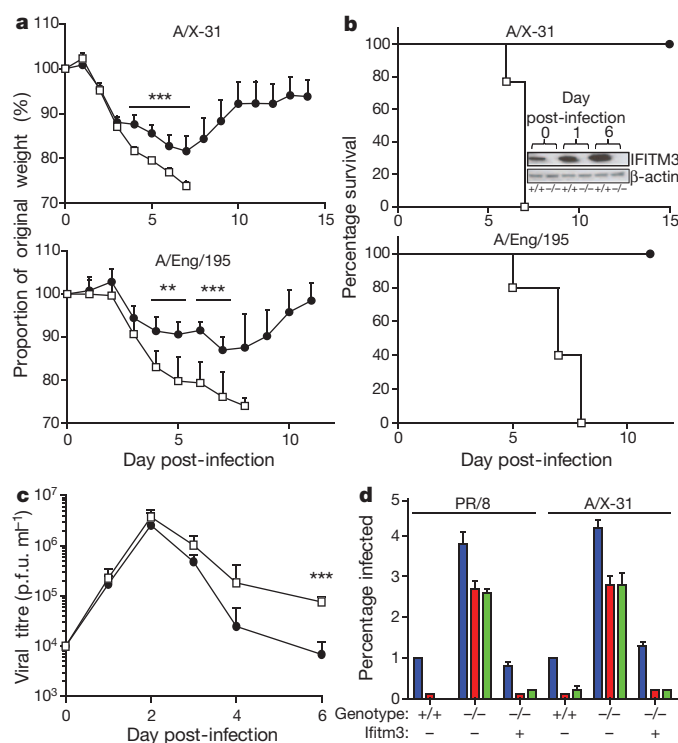


Figure 1 | Influenza A virus replicates to higher levels in *Ifitm3*^{−/−} mice.

a, b, Change in body mass (**a**) and survival (**b**) of wild-type (filled circles) and *Ifitm3*^{−/−} (open squares) mice following intranasal inoculation with A/X-31 and pandemic H1N1/09 Eng/195 influenza ($n > 5$). **b,** Absence of *Ifitm3* expression was verified in the *Ifitm3*^{−/−} mice at all time points, but was seen to increase in wild-type mice. **c,** A/X-31 viral load in the lungs of mice ($n > 4$) was calculated over the course of infection by plaque assay. p.f.u., plaque-forming units. *Ifitm3*^{−/−} murine embryonic fibroblasts ($n = 3$ per condition) stably expressing *Ifitm3* (+), or the empty vector (−) were left untreated (blue), or incubated with IFN- α (red) or IFN- γ (green), then challenged with either A/X-31 or PR/8 influenza. **d,** Twelve hours after infection, the cells were assessed for either haemagglutinin expression (PR/8), or nucleoprotein expression (A/X-31) IFITM3 expression was determined to be present (+) or absent (−) by western blotting (Supplementary Fig. 2). Results show means \pm s.d. Statistical significance was assessed by Student's *t*-test (** $P < 0.01$; *** $P < 0.001$).

¹Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. ²Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Charlestown, Massachusetts 02129, USA. ³Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK. ⁴Division of Genetics and Genomics, The Roslin Institute, University of Edinburgh, Roslin EH25 9RG, UK. ⁵Department of Critical Care Medicine, University of Edinburgh, Edinburgh EH16 4JT, UK. ⁶Institute of Genetics and Biophysics "A. Buzzati-Traverso", National Research Council (CNR), Naples, Italy. ⁷Centre for Respiratory Infection, National Heart and Lung Institute, St Mary's Campus, Imperial College London, W2 1PG, UK. ⁸Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. ⁹Institute of Translational Medicine, University of Liverpool, Alder Hey Children's Hospital, Liverpool L12 2AP, UK. ¹⁰Gastrointestinal Unit, Massachusetts General Hospital, Boston, Massachusetts 02117, USA. ¹¹UCL/MRC Centre for Medical Molecular Virology, Department of Infection, University College London, Cleveland Street, London W1T 4JF, UK.

*Lists of participants and their affiliations appear at the end of the paper.

of infection. However, virus persisted and was not cleared as quickly in *Ifitm3*^{-/-} mice, whose lungs contained tenfold higher levels of replicating virus than the wild-type mice at 6 days post-infection (Fig. 1c). No viral RNA was detected in the heart, brain or spleen of infected wild-type or *Ifitm3*^{-/-} mice over the course of infection, revealing that systemic viraemia was not occurring. Full-genome sequencing of virus removed from the lungs of wild-type and *Ifitm3*^{-/-} mice showed no genetic variation. We demonstrated that IFITM3 protein expression after influenza infection was absent in *Ifitm3*^{-/-} mice but increased substantially in wild-type controls (Fig. 1b and Supplementary Fig. 1). Infection of wild-type and *Ifitm3*^{-/-} mice with a human isolate of pandemic influenza A H1N1 (pH1N1/09) resulted in the same severe pathogenicity phenotype in the *Ifitm3*^{-/-} mice (Fig. 1a, b). Mouse embryonic fibroblast (MEF) lines generated from multiple matched littermates demonstrated that *Ifitm3*^{-/-} cells are infected more readily *in vitro*, and lack much of the protective effects of interferon (IFN). Importantly, the stable restoration of IFITM3 conferred wild-type levels of restriction against either the X-31 strain, or the more pathogenic Puerto Rico/8/34 (PR/8) influenza strain (Fig. 1d and Supplementary Fig. 2). In addition to the role of IFITM3 in restriction of high-pathogenicity H5N1 avian influenza⁷, we also show that it limits infection by recent human influenza A virus isolates and influenza B virus (Supplementary Fig. 3). Therefore, enhanced pathogenesis to diverse influenza viruses is attributable to loss of *Ifitm3* expression and consequential changes in immune defence of the lungs.

Examination of lung pathology showed fulminant viral pneumonia with substantial damage and severe inflammation in the infected *Ifitm3*^{-/-} mice. Lung pathology was characterized by extensive oedema and red blood cell extravasation, as well as pneumonia, haemorrhagic pleural effusion and multiple, large lesions on all lung lobes (Fig. 2a, b and Supplementary Fig. 4). We note that this pathology is similar to that produced by infection of mice and primates with 1918 H1N1 virus⁹⁻¹¹. Given the higher viral load in *Ifitm3*^{-/-} mice and increased replication of influenza A virus in *Ifitm3*-deleted cells *in vitro* (Fig. 1d), we examined both viral nucleic acid and protein distribution in the lung. Influenza virus infection penetrated deeper into the lung tissue in *Ifitm3*^{-/-} compared to wild-type mice whose infection was primarily restricted to the bronchioles, with minimal alveolar infection. Influenza virus was detected throughout the entire lung in *Ifitm3*^{-/-} sections, spreading extensively in both bronchioles and alveoli (Fig. 2c). Histopathology showed marked infiltration of cells and debris into the bronchoalveolar space of *Ifitm3*^{-/-} mice (Fig. 2b and Supplementary Fig. 4b). The extent and mechanism of cell damage was investigated by TdT-mediated dUTP nick end labelling (TUNEL) assay, showing widespread cellular apoptosis occurring 6 days post-infection in *Ifitm3*^{-/-} mice, whereas apoptosis in wild-type lungs was

very limited (Supplementary Fig. 4c). Together, the *Ifitm3*^{-/-} mouse pathology is consistent with infection by high-pathogenicity strains of influenza A virus, where widespread apoptosis occurs by day 6 post-infection, whereas lungs from low-pathogenicity infections were similar to those of wild-type mice, displaying minimal damage^{9,12,13}.

Analysis of cell populations resident in the lung tissue on day 6 post-infection showed that *Ifitm3*^{-/-} mice had significantly reduced proportions of CD4⁺ ($P = 0.004$) and CD8⁺ T cells ($P = 0.02$) and natural killer (NK) cells ($P = 0.0001$), but an elevated proportion of neutrophils ($P = 0.007$) (Fig. 3a). Despite the extensive cellular infiltration (Supplementary Figs 4b, 5a), the absolute numbers of CD4⁺ T-lymphocytes in the lungs of the *Ifitm3*^{-/-} mice were also lower and neutrophils increased compared to wild-type mice (Supplementary Fig. 6). The peripheral blood of infected *Ifitm3*^{-/-} mice showed leukopenia (Supplementary Fig. 5c). Blood differential cell counts indicated marked depletion of lymphocytes on day 2 post-infection in the *Ifitm3*^{-/-} mice ($P = 0.04$) (Fig. 3b), reflecting changes observed previously in high-pathogenicity (but not low-pathogenicity) influenza infections in both humans and animal models^{9,12,14,15}. Heightened cytokine and chemokine levels are also hallmarks of severe influenza infection, having been observed in both human and animal models^{9,16}. We observed exaggerated pro-inflammatory responses in the lungs of *Ifitm3*^{-/-} mice with levels of TNF- α , IL-6, G-CSF and MCP-1 showing the most marked increase (Fig. 3c and Supplementary Fig. 7). This is indicative of the extent of viral spread within the lungs, as TNF- α and IL-6 are released from cells upon infection¹⁷. Consistent with the immunopathology data above, these changes are comparable in level to those seen with non-H5N1 high-pathogenicity influenza infections⁹. Neutrophil chemotaxis, together with elevated proinflammatory cytokine secretion, has previously been reported as one of the primary causes of acute lung injury¹⁸.

To investigate further the extensive damage observed with low-pathogenicity influenza A virus infection in the absence of IFITM3, we infected both wild-type and *Ifitm3*^{-/-} mice with a PR/8 influenza strain deficient for the multi-functional NS1 gene (delNS1)^{19,20}. NS1 is the primary influenza virus interferon antagonist, with multiple inhibitory effects on host immune pathways^{20,21}. We found that delNS1 virus was attenuated in both wild-type and *Ifitm3*^{-/-} mice, and whereas the isogenic PR/8 strain expressing NS1 showed typical high pathogenicity in all mice tested, lower doses of PR/8 influenza (although lethal in both genotypes of mice) caused accelerated weight loss in *Ifitm3*^{-/-} compared to wild-type mice (Supplementary Fig. 8). As delNS1 influenza A virus retains its pathogenicity in IFN-deficient mice¹⁹, this suggests that *Ifitm3*^{-/-} mice can mount an adequate IFN-mediated anti-viral response without extensive morbidity, and that IFITM3 blocking viral replication occurs before NS1-mediated IFN

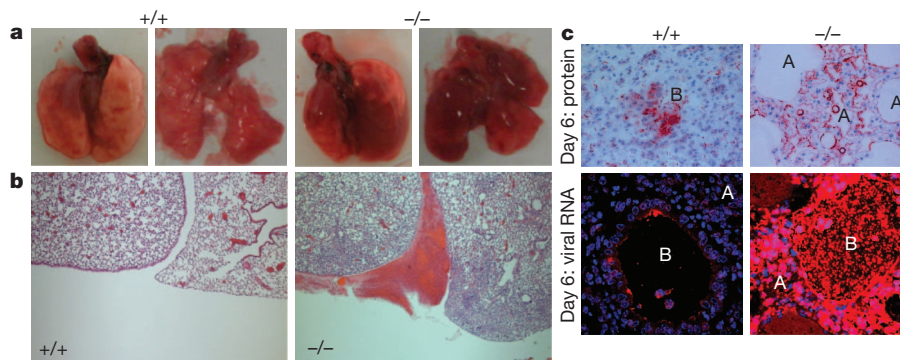


Figure 2 | Pathological examination of infected lungs. a, b, Wild-type mice showed few visible signs of external damage on lung lobes at day 6 post-infection, whereas *Ifitm3*^{-/-} mice showed several large lesions (a, left, ventral view of intact lungs, right, all lobes displayed) resulting from severe oedema and hemorrhagic pleural effusion (b), as well as a markedly higher infiltration of cells and proteinaceous debris into the alveoli and bronchioles. c, Localization

of virus within the lungs on day 6 indicated that virus penetrated deeper and more extensively into the lung tissue in the *Ifitm3*^{-/-} mice, as determined by immunohistochemistry for total influenza protein and detection of virus nucleic acid (virus, red; cell nuclei, blue; A, alveolus; B, bronchiole). Original magnifications were $\times 5$ (b) and $\times 20$ (c).

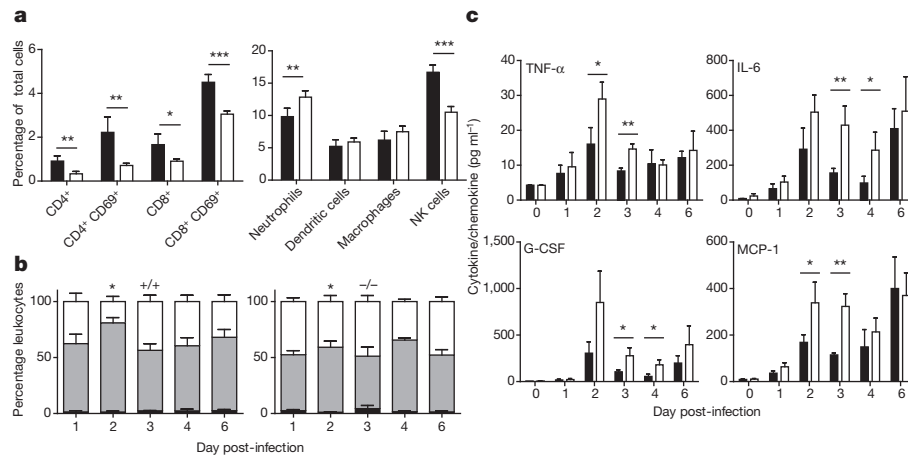


Figure 3 | Altered leukocyte and cytokine response to influenza A infection in *Ifitm3*^{-/-} mice. **a**, Cytometric analysis of proportional resident cell populations in the lungs of mice (+/+, black; -/-, white) showed evidence of lymphopenia in *Ifitm3*^{-/-} mice 6 days post-infection. **b**, Systemic lymphopenia was confirmed through differential analysis of peripheral blood cell counts, which showed a significant depletion of lymphocytes on day 2 post-infection of

antagonism⁷. Therefore, unchecked lung viral replication and an enhanced inflammatory response accounts for the profoundly deleterious effects of viral infection in *Ifitm3*^{-/-} mice.

The human *IFITM3* gene has two exons and is predicted to encode two splice variants that differ by the presence or absence of the first amino-terminal 21 amino acids (Fig. 4a). Currently, 13 non-synonymous,

Ifitm3^{-/-} mice (monocytes, black; lymphocytes, grey; polymorphonuclear leukocytes, white). NK, natural killer. **c**, Levels of pro-inflammatory cytokines were also recorded as being elevated in *Ifitm3*^{-/-} lungs over the course of infection (+/+, black; -/-, white). Results show means ± s.d., *n* = 5. Statistical significance was assessed by Student's *t*-test (**P* < 0.05, ***P* < 0.01, ****P* < 0.001).

13 synonymous, one in-frame stop and one splice site acceptor-altering single nucleotide polymorphisms (SNPs) have been reported in the translated *IFITM3* sequence (Supplementary Table 1). Using tests sensitive to recent positive selection, we can find evidence for positive selection on the *IFITM3* locus in human populations acting over the last tens of thousands of years in Africa (Fig. 4b, c). We therefore

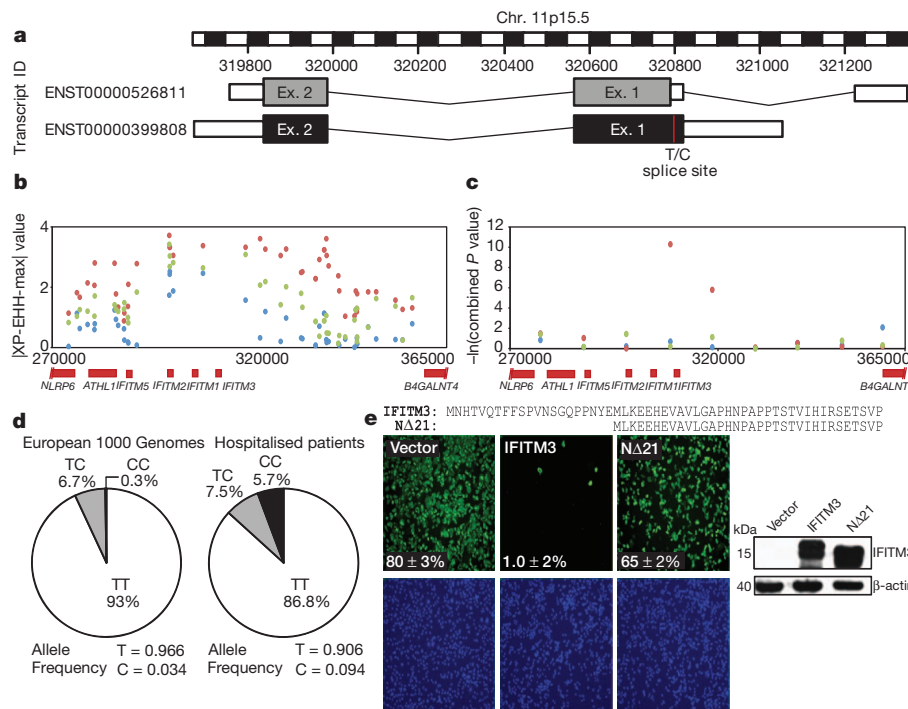


Figure 4 | Single nucleotide polymorphisms of the human *IFITM3* gene. **a**, Multiple single-nucleotide polymorphisms have been identified within the coding region of the human *IFITM3* gene. One such SNP, rs12252 (red), encodes a splice acceptor site altering T/C substitution mutation and may be associated with a truncated protein with an N-terminal 21 amino acid deletion. Therefore two transcripts are predicted to be expressed from the *IFITM3* gene. **b**, **c**, Positive selection analysis using a haplotype-based test, the cross population extended haplotype homozygosity test, maximum value (|XP-EHH-max|, **b**), where data points above 2.7 in the YRI (Africa) (red), 3.9 in the CEU (Europe) (blue) and 5.0 in the CHB+JPT (China and Japan) (green) populations are in the top 1% of values, and using a combination of three allele frequency spectrum-based test statistics (**c**), namely the composite likelihood

ratio (CLR)^{23–25} on 10-kb windows along chromosome 11 encompassing the *IFITM3* locus. Evidence for positive selection is seen only in the YRI. **d**, Mutations recorded through sequencing of patients hospitalized with influenza virus during the H1N1/09 pandemic showed an overrepresentation of individuals with the C allele at SNP rs12252, relative to matched Europeans. **e**, A549 cells transduced to express either full-length (*IFITM3*) or truncated (*Δ21*) *IFITM3* (cell nuclei, blue; virus, green; ×4 magnification) show a reduction in viral restriction when the N-terminal 21 amino acids of *IFITM3* are removed, relative to vector controls (Vector). Alignment of the N termini of full-length (*IFITM3*, top) and truncated *IFITM3* (*Δ21*, bottom). Values represent the mean of the percentage of infected cells ± s.d. (*n* = 3).

sequenced 1.8 kilobases of the *IFITM3* locus encompassing the exons, intron and untranslated regions from 53 individuals who required admission to hospital as a result of pandemic H1N1/09 or seasonal influenza virus infection in 2009–2010. Of these, 86.8% of patients carried majority alleles for all 28 SNPs in the coding sequence of the gene, but 13.2% possessed known variants. In particular, we discovered over-representation in cases of the synonymous SNP rs12252, wherein the majority T allele is substituted for a minority C allele, which alters the first splice acceptor site and may be associated with the *IFITM3* splice variant (ENST00000526811), which encodes an *IFITM3* protein lacking the first 21 amino acids due to the use of an alternative start codon.

The allele frequencies for SNP rs12252 vary in different human populations (Supplementary Table 2). The ancestral (C) allele, reported in chimpanzees, is rare in sub-Saharan African and European populations (derived allele frequency (DAF) 0.093 and 0.026–0.036, respectively), but more frequent in other populations (Supplementary Table 2). SNP rs12252 is notable for its high level of differentiation between Europeans and East Asians, although the fixation index (F_{ST} , a measure of population differentiation) does not reach statistical significance. The genotypes associated with rs12252 in Caucasians hospitalized following influenza infection differ significantly from ethnically matched Europeans in 1000 Genomes sequence data and from genotypes imputed against the June 2011 release of the 1000 Genomes phased haplotypes from the UK, Netherlands and Germany (Wellcome Trust Case Control Consortium 1 (WTCCC1, UK): $P = 0.00006$, Netherlands: $P = 0.00001$, Germany: $P = 0.00007$; Fisher's exact test). Patients' genotypes also depart from Hardy–Weinberg equilibrium ($P = 0.003$), showing an excess of C alleles in this population (Fig. 4d). Principal components analysis of over 100,000 autosomal SNPs showed no evidence of hidden population structure differences between WTCCC controls and a subset of the hospitalised individuals from this study (Supplementary Fig. 9a, b).

To test the functional significance of the *IFITM3* rs12252 polymorphism *in vitro*, we confirmed the genotypes of HapMap lymphoblastoid cell lines (LCLs) homozygous for either the majority (TT) or minority (CC) variant *IFITM3* alleles (Supplementary Fig. 9c). We next challenged the LCLs with influenza A virus and found that the minority (CC) variant was more susceptible to infection, and this vulnerability correlated with lower levels of *IFITM3* protein expression compared to the majority (TT) variant cells (Supplementary Fig. 10). Although we did not detect the *IFITM3* splice variant protein (ENST00000526811) in the CC LCLs, we nonetheless investigated the possible significance of its presence by stably expressing the N-terminally truncated ($\Delta 21$) and wild-type proteins to equivalent levels in human A549 lung carcinoma cell lines before infection with influenza A virus (A/WSN/1933 (WSN/33)). We found that cells expressing the $\Delta 21$ protein failed to restrict viral replication when compared to wild-type *IFITM3* (Fig. 4e), consistent with previous data showing that the amino-terminal 21 amino acids of *IFITM3* are required for attenuation of vesicular stomatitis virus replication *in vitro*⁴. Similar results were obtained using other virulent viral strains (A/California/7/2009 (pH1N1), A/Uruguay/716/2007 (H3N2) and B/Brisbane/60/2008) (Supplementary Fig. 3).

We show here that *IFITM3* expression acts as an essential barrier to influenza A virus infection *in vivo* and *in vitro*. The fulminant viral pneumonia that occurs in the absence of *IFITM3* arises because of uncontrolled virus replication in the lungs, resulting in profound morbidity. In effect, the host's loss of a single immune effector, *IFITM3*, transforms a mild infection into one with remarkable severity. Similarly, the enrichment of the rs12252 C-allele in those hospitalized with influenza infections, together with the decreased *IFITM3* levels and the increased infection of the CC-allele cells *in vitro*, suggests that *IFITM3* also plays a pivotal role in defence against human influenza virus infections. This innate resistance factor is all the more important during encounters with a novel pandemic virus, when the host's acquired

immune defences are less effective. Indeed, *IFITM3*-compromised individuals, and in turn populations with a higher percentage of such individuals, may be more vulnerable to the initial establishment and spread of a virus against which they lack adaptive immunity. In light of its ability to curtail the replication of a broad range of pathogenic viruses *in vitro*, these *in vivo* results suggest that *IFITM3* may also shape the clinical course of additional viral infections in favour of the host, and may have done so over human evolutionary history.

METHODS SUMMARY

Mouse infection. Wild-type and *Ifitm3*^{-/-} mice⁸ (8–10 weeks of age) were intranasally inoculated with 10⁴ p.f.u. of A/X-31 (H3N2) influenza, 200 p.f.u. of A/England/195/09 (pH1N1) influenza, or 50–10³ p.f.u. of A/PR/8/34 (PR/8) or the PR/8 NS1 gene deletion mutant (delNS1)²⁰ (H1N1) in 50 μ l of sterile PBS. Mouse weight was recorded daily as well as monitoring for signs of illness. Mice exceeding 25% total weight loss were killed in accordance with UK Home Office guidelines. Infected lungs were collected on days 1–6 post-infection and quantified for viral load by plaque assay and RT-qPCR with primers to influenza matrix 1 protein.

Pathology of infected *Ifitm3*^{-/-} mice. 5- μ m sections of paraffin-embedded tissue were stained with haematoxylin and eosin and microscopically examined. Apoptosis was assessed by TUNEL using the TACS XL DAB *In Situ* Apoptosis Detection Kit (R&D Systems). Viral RNA was visualized by QuantiGene viewRNA kit (Affymetrix), with a viewRNA probe set designed to the negative stranded vRNA encoding the NP gene of A/X-31 (Affymetrix). Lung tissue was embedded in glycol methacrylate (GMA) and viral antigens stained using M149 polyclonal antibody to influenza A, B (Takara). Single cell suspensions from the lung were characterized by flow cytometry for T-lymphocytes CD4⁺ or CD8⁺, T-lymphocytes (activated) CD4⁺CD69⁺ or CD8⁺CD69⁺, neutrophils CD11b^{hi}CD11c^{Ly6g}⁺, dendritic cells CD11c⁺CD11b^{Ly6g}⁺ MHC class II high, macrophages CD11b⁺CD11c⁺F4/80^{hi}, natural killer cells NKp46⁺CD4⁻CD8⁻.

Sequencing and genetics of human *IFITM3*. The 1.8 kb of human *IFITM3* was amplified and sequenced to identify single nucleotide polymorphisms (SNPs). SNP rs12252 was identified and compared to allele and genotype frequencies from 1000 Genomes sequencing data from different populations including 1000 Genomes imputed. SNP rs12252 allele frequencies were determined in the publicly available genotype data sets of WTCCC1 ($n = 2,938$) and previously published data sets genotyped from the Netherlands ($n = 8,892$) and Germany ($n = 6,253$)²².

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 September 2011; accepted 8 February 2012.

Published online 25 March 2012.

- Brass, A. L. *et al.* The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. *Cell* **139**, 1243–1254 (2009).
- Jiang, D. *et al.* Identification of five interferon-induced cellular proteins that inhibit West Nile virus and dengue virus infections. *J. Virol.* **84**, 8332–8341 (2010).
- Yount, J. S. *et al.* Palmitoylation profiling reveals S-palmitoylation-dependent antiviral activity of IFITM3. *Nature Chem. Biol.* **6**, 610–614 (2010).
- Weidner, J. M. *et al.* Interferon-induced cell membrane proteins, IFITM3 and tetherin, inhibit vesicular stomatitis virus infection via distinct mechanisms. *J. Virol.* **84**, 12646–12657 (2010).
- Huang, I. C. *et al.* Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. *PLoS Pathog.* **7**, e1001258 (2011).
- Schoggins, J. W. *et al.* A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature* **472**, 481–485 (2011).
- Feeley, E. M. *et al.* IFITM3 inhibits influenza A virus infection by preventing cytosolic entry. *PLoS Pathog.* **7**, e1002337 (2011).
- Lange, U. C. *et al.* Normal germ line establishment in mice carrying a deletion of the *Ifitm/Fragilis* gene family cluster. *Mol. Cell. Biol.* **28**, 4688–4696 (2008).
- Belser, J. A. *et al.* Pathogenesis of pandemic influenza A (H1N1) and triple-reassortant swine influenza A (H1) viruses in mice. *J. Virol.* **84**, 4194–4203 (2010).
- Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77–80 (2005).
- Kobasa, D. *et al.* Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus. *Nature* **445**, 319–323 (2007).
- Tumpey, T. M., Lu, X. H., Morken, T., Zaki, S. R. & Katz, J. M. Depletion of lymphocytes and diminished cytokine production in mice infected with a highly virulent influenza A (H5N1) virus isolated from humans. *J. Virol.* **74**, 6105–6116 (2000).
- Kobasa, D. *et al.* Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus. *Nature* **431**, 703–707 (2004).
- Maines, T. R. *et al.* Pathogenesis of emerging avian influenza viruses in mammals and the host innate immune response. *Immunol. Rev.* **225**, 68–84 (2008).
- Perrone, L. A., Plowden, J. K., Garcia-Sastre, A., Katz, J. M. & Tumpey, T. M. H5N1 and 1918 pandemic influenza virus infection results in early and excessive infiltration of macrophages and neutrophils in the lungs of mice. *PLoS Pathog.* **4**, e1000115 (2008).

16. Fukuyama, S. & Kawaoka, Y. The pathogenesis of influenza virus infections: the contributions of virus and host factors. *Curr. Opin. Immunol.* **23**, 481–486 (2011).
17. Julkunen, I. *et al.* Inflammatory responses in influenza A virus infection. *Vaccine* **19**, S32–S37 (2000).
18. Yum, H. K. *et al.* Involvement of phosphoinositide 3-kinases in neutrophil activation and the development of acute lung injury. *J. Immunol.* **167**, 6601–6608 (2001).
19. Garcia-Sastre, A. *et al.* Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **322**, 324–330 (1998).
20. Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* **89**, 2359–2376 (2008).
21. Billharz, R. *et al.* The NS1 protein of the 1918 pandemic influenza virus blocks host interferon and lipid metabolism pathways. *J. Virol.* **83**, 10557–10570 (2009).
22. Anttila, V. *et al.* Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nature Genet.* **42**, 869–873 (2011).
23. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
24. Fay, J. C. & Wu, C.-I. Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405–1413 (2000).
25. Nielsen, R. *et al.* Genomic scans for selective sweeps using SNP data. *Genome Res.* **15**, 1566–1575 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank C. Brandt for maintaining mouse colony health and well-being and T. Hussell for provision of A/X-31 virus. We also thank D. Gurdasani and M. Sandhu for statistical analysis of genotype frequencies. We also thank M. Hu and I. Gallego Romero for calculating genome-wide selection statistics. This work was supported by the Wellcome Trust. The MOSAIC work was supported by Imperial's Comprehensive Biomedical Research Centre (cBRC), the Wellcome Trust (090382/Z/09/Z) and Medical Research Council UK. The GenISIS work was supported by the Chief Scientist Office (Scotland). A.L.B. is the recipient of a Charles H. Hood Foundation Child Health Research Award, and is supported by grants from the Phillip T. and Susan M. Ragon Institute Foundation, the Bill and Melinda Gates Foundation's Global Health Program and the National Institute of Allergy and Infectious Diseases (R01AI091786). J.K.B. is supported by a Wellcome Trust Clinical Lectureship (090385/Z/09/Z) through the Edinburgh Clinical Academic Track (ECAT). We acknowledge the assistance of K. Alshafi, E. Bailey, A. Bermingham, M. Berry, C. Bloom, E. Brannigan, S. Bremang, J. Clark, M. C. Cox, M. Cross, L. A. Cumming, S. Dyas, J. England-Smith, J. Enstone, D. Ferreira, N. Goddard, A. Godlee, S. Gormley, M. Guiver, M. O. Hassan-Ibrahim, H. Hill, P. Holloway, K. Hoschler, G. Houghton, F. Hughes, R. R. Israel, A. Jepson, K. D. Jones, W. P. Kelleher, M. Kidd, K. Knox, A. Lackenby, G. Lloyd, H. Longworth, M. Minns, S. Moorkerjee, S. Mt-Isa, D. Muir, A. Paras, V. Pascual, L. Rae, S. Rodenhurst, F. Rozakeas, E. Scott, E. Sergi, N. Shah, V. Sutton, J. Vernazza, A. W. Walker, C. Wenden, T. Wotherspoon, A. D. Wright, F. Wurie and the clinical and laboratory staff of the Alder Hey Children's NHS Foundation Trust, Brighton & Sussex University Hospitals NHS Trust, Central Manchester University Hospitals NHS Foundation Trust, Chelsea and Westminster Hospital NHS Foundation Trust, Alder Hey Children's Hospital and Liverpool School of Tropical Medicine, Health Protection Agency Microbiology Services Colindale, Imperial College Healthcare NHS Trust, Liverpool Women's NHS Foundation Trust, Royal Liverpool and Broadgreen University Hospitals NHS Trust, Royal Brompton and Harefield NHS Foundation Trust, The Roslin Institute, Edinburgh, University Hospitals Coventry and Warwickshire NHS Trust. The MOSAIC consortium was supported by several Comprehensive Local Research Networks (CLRNs), the National Institute for Health Research (NIHR), UK, and by the Biomedical Research Centre (BRC) and Unit (BRU) funds. Finally, we thank all patients and their relatives for their generous agreement to inclusion in this study.

Author Contributions A.R.E., G.D., A.L.B. and P.K. designed the study; A.R.E., P.J.O., G.D., A.L.B. and P.K. wrote the manuscript; A.R.E. performed experiments and analysed data; S.C. designed experiments and performed all live animal work; S.E.S. sequenced and analysed the human *Ifitm3* gene; R.S.W., S.E.S., C.R.C., J.S.S., S.P.J., T.P., E.M.F., A.L.B. and L.K. performed experiments; D.J.A. created the genetically-modified *Ifitm3*^{-/-} mouse line; H.M.W. and P.D. made the influenza virus strains and advised on virology; D.G. performed microscopy; Y.X., V.C. and C.T.-S. performed positive selection analyses; V.A. and A.P. performed imputation and analysis of 1000 Genomes data; E.M.F., C.R.C. and A.L.B. performed *in vitro* viewRNA experiments; recruitment and selection of hospitalised individuals infected with influenza virus was co-ordinated by J.K.B., D.A.H. and T.S.W. (GenISIS) and R.L.S., S.B.G., J.D., J.K.B., D.A.H. and P.J.O. (MOSAIC).

Author Information *IFITM3* sequences are deposited in GenBank under the accession numbers JQ610570–JQ610621. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.K. (pk5@sanger.ac.uk) or A.L.B. (abrass@partners.org).

MOSAIC Core Investigators

Benaroya Research Institute, USA D. Chaussabel¹; **Gartnavel General Hospital, Greater Glasgow, UK** W. E. Adamson², W. F. Carman²; **Health Protection Agency, UK** C. Thompson³, M. C. Zambon³; **Imperial College London, UK** P. Aylin⁴, D. Ashby⁴, W. S. Barclay⁴, S. J. Brett⁴, W. O. Cookson⁴, L. N. Drumright⁴, J. Dunning⁴, R. A. Elderfield⁴, L. Garcia-Alvarez⁴, B. G. Gazzard⁴, M. J. Griffiths⁴, M. S. Habibi⁴, T. T. Hansel⁴, J. A. Herberg⁴, A. H. Holmes⁴, T. Hussell⁴, S. L. Johnston⁴, O. M. Kon⁴, M. Levin⁴, M. F. Moffatt⁴, S. Nadel⁴,

P. J. Openshaw⁴, J. O. Warner⁴; **Liverpool School of Tropical Medicine, UK** S. J. Aston⁵, S. B. Gordon⁵; **National Institute for Medical Research, UK** A. Hay⁶, J. McCauley⁶, A. O'Garra⁶; **Roche, Nutley, USA** J. Banachereau⁷; **University College London, UK** A. Hayward⁸, P. Kellam⁸; **University of Edinburgh, UK** J. K. Baillie⁹, D. A. Hume⁹, P. Simmons⁹; **University of Liverpool, UK** P. S. McNamara¹⁰, M. G. Semple¹⁰, R. L. Smyth¹⁰; **University of Nottingham, UK** J. S. Nguyen-Van-Tam¹¹; **University of Oxford, UK** L.-P. Ho¹², A. J. McMichael¹²; **Wellcome Trust Sanger Institute, UK** P. Kellam¹³.

GenISIS Investigators

Critical Care Medicine, University of Edinburgh K. Everingham¹⁴, H. Dawson¹⁴, D. Hope¹⁴, P. Ramsay¹⁴, T. S. Walsh (Local Lead Investigator)¹⁴; **Generation Scotland, University of Edinburgh Molecular Medicine Centre** A. Campbell¹⁵, S. Kerr¹⁵; **Intensive Care National Audit & Research Centre, London** D. Harrison¹⁶, K. Rowan¹⁶; **Intensive Care Unit, Aberdeen Royal Infirmary** J. Addison¹⁷, N. Donald¹⁷, S. Galt¹⁷, D. Noble¹⁷, J. Taylor¹⁷, N. Webster (Local Lead Investigator)¹⁷; **Intensive Care Unit, Ayr Hospital** I. Taylor (Local Lead Investigator)¹⁸; **Intensive Care Unit, Borders General Hospital, Melrose** J. Aldridge (Local Lead Investigator)¹⁹, R. Dornan¹⁹, C. Richard¹⁹; **Intensive Care Unit, Crosshouse Hospital, Kilmarnock** D. Gilmour²⁰, R. Simmons (Local Lead Investigator)²⁰, R. White (Local Lead Investigator)²⁰; **Intensive Care Unit, Dumfries and Galloway Royal Infirmary** C. Jardine²¹, D. Williams (Local Lead Investigator)²¹; **Intensive Care Unit, Glasgow Royal Infirmary** M. Booth (Local Lead Investigator)²², T. Quasim²²; **Intensive Care Unit, Hairmyres Hospital, Lanarkshire** V. Watson²³; **Intensive Care Unit, Inverclyde Royal Hospital, Greenock** P. Henry²⁴, F. Munro²⁴; **Intensive Care Unit, Monklands Hospital, Airdrie** L. Bell²⁵, J. Ruddy (Local Lead Investigator)²⁵; **Intensive Care Unit, Ninewells Hospital, Dundee** S. Cole (Local Lead Investigator)²⁶, J. Southward²⁶; **Intensive Care Unit, Queen Margaret Hospital, Dunfermline** P. Alcoat²⁷, S. Gray²⁷, M. McDougall (Local Lead Investigator)²⁷; **Intensive Care Unit, Raigmore Hospital, Inverness** J. Matheson²⁸, J. Whiteside (Local Lead Investigator)²⁸; **Intensive Care Unit, Royal Alexandra Hospital, Paisley** D. Alcorn²⁹, K. Rooney (Local Lead Investigator)²⁹, R. Sundaram²⁹; **Intensive Care Unit, Southern General Hospital, Glasgow** G. Imrie (Local Lead Investigator)³⁰; **Intensive Care Unit, St John's Hospital, Livingston** J. Bruce³¹, K. McGuigan³¹, S. Moultrie (Local Lead Investigator)³¹; **Intensive Care Unit, Stirling Royal Infirmary** C. Cairns (Local Lead Investigator)³², J. Grant³², M. Hughes³²; **Intensive Care Unit, Stobhill Hospital, Glasgow** C. Murdoch (Local Lead Investigator)³³; **Intensive Care Unit, Victoria Hospital, Glasgow** A. Davidson (Local Lead Investigator)³⁴; **Intensive Care Unit, Western General Hospital, Edinburgh** G. Harris³⁵, R. Paterson³⁵, C. Wallis (Local Lead Investigator)³⁵; **Intensive Care Unit, Western Infirmary, Glasgow** S. Binning (Local Lead Investigator)³⁶, M. Pollock³⁶; **Wellcome Trust Clinical Research Facility, Edinburgh** J. Antonelli³⁷, A. Duncan³⁷, J. Gibson³⁷, C. McCulloch³⁷, L. Murphy³⁷; **Roslin Institute, University of Edinburgh** C. Haley³⁸, G. Faulkner³⁸, T. Freeman³⁸, D. A. Hume³⁸ & J. K. Baillie (Principal Investigator)³⁸.

¹Benaroya Research Institute, 1201 9th Avenue, Seattle, Washington 98101-2795, USA.

²West of Scotland Specialist Virology Centre, Gartnavel General Hospital, Glasgow and Clyde Health Board, 1053 Great Western Road, Glasgow G12 0YN, UK. ³Health Protection Agency, Microbiology Services Colindale, 61 Colindale Avenue, London NW9 5EQ, UK.

⁴Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, UK.

⁵Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. ⁶National Institute for Medical Research (NIMR), The Ridgeway, London NW7 1AA, UK. ⁷Roche, 340 Kingsland Street, Nutley, New Jersey 07110-1199, USA. ⁸University College London, Gower Street, London, WC1E 6BT, UK. ⁹The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK. ¹⁰Department of Women's and Children's Health, Institute of Translational Medicine, University of Liverpool, Alder Hey Children's Hospital, Liverpool L12 2AP, UK. ¹¹Epidemiology and Public Health, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham NG5 1PB, UK. ¹²The Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK. ¹³Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. ¹⁴Department of Critical Care Medicine, The Queen's Medical Research Institute, College of Medicine and Veterinary Medicine, University of Edinburgh, 47 Little France Drive, Edinburgh EH16 4JT, UK. ¹⁵Generation Scotland, University of Edinburgh, Molecular Medicine Centre, Western General Hospital, Crewe Road South, Edinburgh EH4 2XU, UK. ¹⁶Intensive Care National Audit & Research Centre, Tavistock House, Tavistock Square, London WC1H 9HR, UK. ¹⁷Intensive Care Unit, Aberdeen Royal Infirmary, Foresterhill, Aberdeen AB25 2ZN, UK. ¹⁸Intensive Care Unit, Ayr Hospital, Dailmellington Road, Ayr KA6 6DX, UK. ¹⁹Intensive Care Unit, Borders General Hospital, Melrose TD6 9BS, UK. ²⁰Intensive Care Unit, Crosshouse Hospital, Kilmarnock KA2 0BE, UK. ²¹Intensive Care Unit, Dumfries and Galloway Royal Infirmary, Dumfries DG1 4AP, UK. ²²Intensive Care Unit, Glasgow Royal Infirmary, Glasgow G4 0SF, UK. ²³Intensive Care Unit, Hairmyres Hospital, Lanarkshire G75 8RG, UK. ²⁴Intensive Care Unit, Inverclyde Royal Hospital, Greenock PA1 6OXN, UK. ²⁵Intensive Care Unit, Monklands Hospital, Airdrie ML6 0JS, UK. ²⁶Intensive Care Unit, Ninewells Hospital, Dundee DD1 9SY, UK. ²⁷Intensive Care Unit, Queen Margaret Hospital, Dunfermline KY12 0SU, UK. ²⁸Intensive Care Unit, Raigmore Hospital, Inverness IV2 3UJ, UK. ²⁹Intensive Care Unit, Royal Alexandra Hospital, Paisley PA2 9PN, UK. ³⁰Intensive Care Unit, Southern General Hospital, Glasgow G51 4TF, UK. ³¹Intensive Care Unit, St John's Hospital, Livingston EH54 6PP, UK. ³²Intensive Care Unit, Stirling Royal Infirmary, Stirling FK8 2AU, UK. ³³Intensive Care Unit, Stobhill Hospital, Glasgow G21 3UW, UK. ³⁴Intensive Care Unit, Victoria Hospital, Glasgow G42 9TY, UK. ³⁵Intensive Care Unit, Western General Hospital, Crewe Road South, Edinburgh EH4 2XU, UK. ³⁶Intensive Care Unit, Western Infirmary, Glasgow G11 6NT, UK. ³⁷Wellcome Trust Clinical Research Facility, Western General Hospital, Crewe Road South, Edinburgh EH4 2XU, UK. ³⁸Division of Genetics and Genomics, The Roslin Institute, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK.

METHODS

Mouse infection. Background-matched wild-type (>95% C57BL/6) and *Ifitm3*^{-/-} mice⁸ 8–10 weeks of age were maintained in accordance with UK Home Office regulations, UK Animals Scientific Procedures Act 1986 under the project licence PPL80/2099. This licence was reviewed by The Wellcome Trust Sanger Institute Ethical Review Committee. Groups of >5 isoflurane-anaesthetized mice of both genotype were intranasally inoculated with 10⁴ p.f.u. of A/X-31 influenza in 50 µl of sterile PBS. In some experiments A/X-31 was substituted with 200 p.f.u. of A/England/195/09 influenza, or 50–10³ p.f.u. of A/PR/8/34 (PR/8) or an otherwise isogenic virus with a deletion of the *NSI* gene (delNSI)¹⁹, made as described²⁶. Their weight was recorded daily and they were monitored for signs of illness. Mice exceeding 25% total weight loss were killed in accordance with UK Home Office guidelines. Littermate controls were used in all experiments.

Influenza virus quantification. Lungs from five mice per genotype were collected on days 1, 2, 3, 4 and 6 post-infection, weighed and homogenized in 5% weight/volume (w/v) of Leibovitz's L-15 medium (Invitrogen) containing antibiotic-antimycotic (Invitrogen). Samples were quantified for viral load by plaque assay in tenfold serial dilutions on Madin–Darby canine kidney (MDCK) cell monolayers overlaid with 1% Avicell medium²⁷. Lungs were subjected to two freeze-thaw cycles before titration. Virus was also quantified by quantitative PCR with reverse transcription (qRT–PCR), wherein RNA was first extracted from lung, heart, brain and spleen using the RNeasy Mini Plus Kit (Qiagen). Purified RNA was normalized by mass and quantified with SYBR Green (Qiagen) using the manufacturer's instructions and 0.5 µM primers for influenza matrix 1 protein (*M1*) forward: 5'-TGA GTCTTCTAACCGAGGTC-3', reverse: 5'-GGTCTTGTCTTTAGCCATTCC-3' (Sigma-Aldrich) and mouse β-actin (*Actb*) forward: 5'-CTAAGGCCAACCGTG AAAAG-3', reverse: 5'-ACCAGAGGCATACAGGGACA-3'. qPCR was performed on a StepOnePlus machine (Applied Biosystems) and analysed with StepOne software v2.1 (Applied Biosystems).

Western blotting. Lungs were homogenized in 5% w/v of Tissue Protein Extraction Reagent (Thermo Scientific) containing cComplete Protease Inhibitor (Roche). Total protein was quantified by BCA assay (Thermo Scientific) and was normalized before loading into wells. Proteins were visualized with the following indicated primary antibodies: anti-mouse IFITM2 rabbit polyclonal was purchased from Santa Cruz Biotechnology (catalogue no. sc-66828); anti-Fragilis (*Ifitm3*) rabbit polyclonal antibody was from Abcam (catalogue no. ab15592). The IFITM3 and NA21 western blot using the A549 stable cell lines were probed with the anti-IFITM1 antibody from Prosci (catalogue no. 5807), which recognizes a conserved portion of the IFITM1, IFITM2 and IFITM3 proteins which is still present even in the absence of the first twenty one N-terminal amino acids. The LCL blots (including the A549 cell line lysate controls) were probed with either an antibody which is specific for the N terminus of IFITM3 (rabbit anti-IFITM3 (N-terminal amino acids 8–38) (Abgent, catalogue no. AP1153a)), or with anti-IFITM1 antibody from Prosci (catalogue no. 5807), as well as rabbit anti-MX1 (Proteintech, catalogue no. 13750-1-AP) and mouse anti-GAPDH (clone GAPDH-71.1) (Sigma, catalogue no. G8795). For the LCL immunoblots all antibodies were diluted in DPBS (Sigma) containing 0.1% Tween 20 (Sigma) and 5% non-fat dried milk (Carnation) and incubated overnight at 4 °C. All primary antibodies were consequently bound to the corresponding species-appropriate horseradish peroxidase-conjugated secondary antibodies (Dako). Actin antibody was purchased from either Abcam or Sigma, mouse monoclonal, catalogue no. A5316.

Pathological examination. 5-µm sections of paraffin-embedded tissue were stained with haematoxylin and eosin (Sigma-Aldrich) and were examined and scored twice, once by a pathologist under blinded conditions. The TUNEL assay for apoptosis was conducted using the TACS XL DAB *In Situ* Apoptosis Detection Kit (R&D Systems).

Immunofluorescent tissue staining: protein. Lung tissue was embedded in glycol methacrylate (GMA) to visualize the spread of viral protein, as described previously²⁸. Briefly, 2-µm sections were blocked with 0.1% sodium azide and 30% hydrogen peroxide followed by a second block of RPMI 1640 (Invitrogen) containing 10% fetal calf serum (Sigma-Aldrich) and 1% bovine serum albumin (Invitrogen). Viral antigen was stained using M149 polyclonal antibody to influenza A, B (Takara) and visualized with a secondary goat anti-rabbit antibody conjugated to alkaline phosphatase (Dako). Sections were counterstained with haematoxylin (Sigma-Aldrich). Murine IFITM1 and IFITM3 protein expression in lung sections from either uninfected mice, or those 2 days post-infection with A/X-31, were immunostained with either anti-IFITM1 antibody (Abcam, catalogue no. ab106265) or anti-fragilis (anti-*Ifitm3*) rabbit polyclonal antisera (Abcam, catalogue no. ab15592). Sections were also stained for DNA with Hoechst 33342 (Sigma).

Immunofluorescent staining: RNA. Viral RNA was visualized in 5-µm paraffin-embedded sections using the QuantiGene viewRNA kit (Affymetrix). Briefly, sections were rehydrated and incubated with proteinase K. They were subsequently

incubated with a viewRNA probe set designed against the negative stranded viral RNA encoding the NP gene of A/X-31 (Affymetrix). The signal was amplified before incubation with labelled probes and visualized.

Flow cytometry. Single-cell suspensions were generated by passing lungs twice through a 100-µm filter before lysing red blood cells with RBC lysis buffer (eBioscience) and assessing for cell viability via Trypan blue exclusion. Cells were characterized by flow cytometry as follows: T-lymphocytes CD4⁺ or CD8⁺, T-lymphocytes (activated) CD4⁺CD69⁺ or CD8⁺CD69⁺, neutrophils CD11b^{hi}CD11c⁻Ly6g⁺, dendritic cells CD11c⁺CD11b^{lo}Ly6g^{lo} MHC class II high, macrophages CD11b⁺CD11c⁺F4/80^{hi}, natural killer cells NKp46⁺CD4⁻CD8⁻. All antibodies (Supplementary Table 3) were from BD Bioscience, except CD69 and F4/80, which were from AbD Serotec. Samples were run on a FACSARIA II (BD Bioscience) and visualized using FlowJo 7.2.4. Data were analysed statistically and graphed using Prism 5.0 (GraphPad Software).

Peripheral leukocyte analysis. Mice (*n* = 3 per genotype per day) were bled on days 0, 1, 2, 3, 4 and 6 by tail vein puncture. Leukocyte counts were determined by haemocytometer, whereas blood cell differential counts were calculated by counting from duplicate blood smears stained with Wright–Giemsa stain (Sigma-Aldrich). At least 100 leukocytes were counted per smear. All blood analyses were conducted in a blinded fashion. Data were analysed statistically and graphed using Prism 5.0 (GraphPad Software).

Cytokine/chemokine analysis. Lungs were collected and homogenized on days 0, 1, 2, 3, 4 and 6 post-infection from four mice of each genotype. G-CSF, GM-CSF, IFN-γ, IL-10, IL-1α, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-9, IP-10, KC-like, MCP-1, MIP-1α, RANTES and TNF-α were analysed using a mouse antibody bead kit (Millipore) according to the manufacturer's instructions on a Luminex FlexMAP3D. Results were analysed and quality control checked using Masterplex QT 2010 and Masterplex Readerfit 2010 (MiraiBio). Data were analysed statistically and graphed using Prism 5.0 (GraphPad Software).

Murine embryonic fibroblast generation, transduction and infectivity assays.

Adult *Ifitm3*^{-/-} mice⁸ were intercrossed and fibroblasts (MEFs) were derived from embryos at day 13.5 of gestation, as described previously¹. MEFs were genotyped by PCR (Thermo-Start Taq DNA Polymerase, ABgene) on embryo tail genomic DNA using primers and the cycle profile described previously⁸ to detect the presence of the wild-type allele (450 base pairs band) and the targeted/knock-out allele (650 bp band). MEFs were cultured in DMEM containing 10% FBS, 1× MEM essential amino acids, 1× 2-mercapto-ethanol (Gibco). MEFs were transduced with vesicular stomatitis virus G (VSV-G) pseudotyped retroviruses expressing either the empty vector control (pQXCIP, Clontech), or one expressing *Ifitm3*, as previous described¹. After puromycin selection the respective cell lines were challenged with either A/X-31 virus (multiplicity of infection (m.o.i.) 0.3–0.4) or PR/8 (m.o.i. 0.4). For PR/8 infections, after 12 h the media was removed and the cells were then fixed with 4% formalin and stained with purified anti-haemagglutinin monoclonal antibody (Hybridoma HA36-4-5.2, Wistar Institute). For A/X-31 experiments, cells were processed comparably as above, but in addition were permeabilized, followed by immunostaining for NP expression (NP (clone H16-L10-4R5) mouse monoclonal (Millipore MAB8800)). Both sets of experiments were completed using an Alexa Fluor 488 goat anti-mouse secondary antibody at 1:1,000 (A11001, Invitrogen). The cells were imaged on an automated Image Express Micro microscope (Molecular Devices), and images were analysed using the MetaMorph Cell Scoring software program (Molecular Devices). Cytokines: cells were incubated with cytokines for 24 h before viral infection. Murine interferon α (PBL Interferon Source, catalogue no. 12100-1) and IFN-γ (PBL Interferon Source, catalogue no. 12500-2) were used at 500–2,500 U ml⁻¹, and 100–300 ng ml⁻¹, respectively.

A549 transduction and infectivity assays. A549 cells (ATCC catalogue no. CCL-185) were grown in complete media (DMEM (Invitrogen catalogue no. 11965) with 10% FBS (Invitrogen)). A549 stable cell lines were made by gamma-retroviral transduction using either the empty vector control virus (pQXCIP, Clontech), the full-length human *IFITM3* complementary DNA, or a truncated human *IFITM3* cDNA which is missing the first 21 amino acids (NA21). After puromycin selection, expression of the IFITM3 and NA21 proteins were confirmed by western blotting using an 18% SDS–PAGE gel and an anti-IFITM3 antibody that was raised against the conserved intracellular loop (CIL) of IFITM3 (Proteintech). A549 cell lines were challenged with one of the following strains: A/WSN/33 (a gift of P. Palese), A/California/7/2009, A/Uruguay/716/2007 and B/Brisbane/60/2008 (gift of J. Malbry) for 12 h, then fixed with 4% paraformaldehyde (PFA) and immunostained with anti-HA antibody (Wistar collection) or anti-NP antibodies (Abcam), or Millipore clone H16-L10-4R5 anti-influenza A virus antibody). Percent infection was calculated from immunofluorescent images as described for the MEF experiments above. Alternatively, cells were transduced with lentiviral vectors to express green fluorescent protein (GFP) or IFITM3 and were stained with anti-NP antibody (Abcam) and analysed by flow cytometry following

challenge with B/Bangladesh/3333/2007 virus (NIMR, England). For the immunofluorescence-based viral titring experiments, virus-containing supernatant was collected from the indicated A549 cell line cultures after 12 h of infection with WSN/33 (part one). Next this supernatant was used to infect MDCK cells (ATCC) in a well by well manner (part two). Both the A549 and MDCK cells were then processed to detect viral HA expression as described above.

LCL infectivity assays. LCL TT and LCL CC cells were grown in RPMI-1640 (Sigma-Aldrich) containing 10% FCS, 2 mM L-glutamine, 1 mM sodium pyruvate, 1× MEM non-essential amino acids solution, and 20 mM HEPES (all from Invitrogen). For infectivity assays, LCL cells were either treated with recombinant human IFN- α 2 (PBL Interferon Source, catalogue no. 11100) at 100 units per ml or DPBS (Sigma-Aldrich) for 16 h. The LCL cells were then counted, resuspended at a concentration of 5×10^5 cells per ml, and plated on a 96-well round-bottom plate (200 μ l cell suspension per well). The cells were then challenged with WSN/33 influenza A virus (m.o.i. 0.1). After 18 h, the cells were washed twice with 250 μ l MACS buffer (DPBS containing 2% FCS and 2 mM EDTA (Sigma-Aldrich)). The cells were fixed and permeabilized using the BD Cytotfix/Cytoperm Fixation/Permeabilization Kit (BD Biosciences), following the manufacturer's instructions. Briefly, the cells were resuspended in 100 μ l of Cytotfix/Cytoperm Fixation and Permeabilization solution and incubated at 4 °C for 20 min. The cells were then washed twice with 250 μ l 1× Perm/Wash buffer and resuspended in 50 ml 1× Perm/Wash buffer containing a 2 μ g ml⁻¹ solution of a fluorescein isothiocyanate (FITC)-conjugated mouse monoclonal antibody against influenza A virus NP (clone 431, Abcam, catalogue no. ab20921). The cells were incubated in the diluted antibody solution for 1 h at 4 °C, washed twice with 250 μ l 1× Perm/Wash buffer, resuspended in 200 μ l MACS buffer, and analysed by flow cytometry using a BD FACS Calibur (BD Biosciences).

Ethics and sampling. We recruited patients with confirmed seasonal influenza A or B virus or pandemic influenza A pH1N1/09 infection who required hospitalization in England and Scotland between November 2009 and February 2011. Patients with significant risk factors for severe disease and patients whose daily activity was limited by co-morbid illness were excluded. 53 patients, 29 male and 24 female, average age 37 (range 2–62) were selected. 46 (88%) had no concurrent co-morbidities. The remaining 6 had the following comorbid conditions: hypertension (3 patients), alcohol dependency and cerebrovascular disease (1 patient), bipolar disorder (1 patient) and kyphoscoliosis (1 patient). Four patients were pregnant. Where assessed, 36 patients had normal body mass (69%), one had a body mass index <18.5 and 10 had a body mass index between 25 and 39.9 and one a body mass index >40. Seasonal influenza A H3N2, influenza B and pandemic influenza A pH1N1/09 were confirmed locally by viral PCR or serological tests according to regional protocols. Consistent with the prevalent influenza viruses circulating in the UK between 2009 and 2011 (ref. 29) 44 (85%) had pH1N1/09, 2 had pH1N1/09 and influenza B co-infection, 4 had influenza B and 2 had non-subtyped influenza A virus infection. Of the adults, 24 required admission to an intensive care unit (ICU) and 1 required admission to a high dependency unit (HDU). The remainder were managed on medical wards and survived their illnesses. The GenISIS study was approved by the Scotland 'A' Research Ethics Committee (09/MRE00/77) and the MOSAIC study was approved by the NHS National Research Ethics Service, Outer West London REC (09/H0709/52, 09/MRE00/67).

Consent was obtained directly from competent patients, and from relatives/friends/welfare attorneys of incapacitated patients. Anonymized 9-ml EDTA blood samples were transported at ambient temperature. DNA was extracted using a Nucleon Kit (GenProbe) with the BACC3 protocol. DNA samples were resuspended in 1 ml TE buffer pH 7.5 (10 mM Tris-Cl pH 7.5, 1 mM EDTA pH 8.0).

Sequencing and genetics. Human *IFITM3* sequences were amplified from DNA obtained from peripheral blood by nested PCR (GenBank accession numbers JQ610570 to JQ610621). The first round used primers forward: 5'-TGAGGGT TATGGGAGACGGGGT-3' and reverse: 5'-TGCTCACGGCAGGAGGCC-3', followed by an additional round using primers forward: 5'-GCTTTGGGGGA ACGGTTGTG-3' and reverse: 5'-TGCTCACGGCAGGAGGCCGA-3'. The

1.8-kb *IFITM3* band was gel-extracted and purified using the QIAquick Gel Extraction Kit (Qiagen). *IFITM3* was Sanger-sequenced on an Applied Biosystems 3730xl DNA Analyzer (GATC Biotech) using a combination of eight sequencing primers (Supplementary Table 4). Single-nucleotide polymorphisms were identified by assembly to the human *IFITM3* encoding reference sequence (accession number NC_000011.9) using Lasergene (DNAStar). Homozygotes were called based on high, single base peaks with high Phred quality scores, whereas heterozygotes were identified based on low, overlapping peaks of two bases with lower Phred quality scores relative to surrounding base calls (Supplementary Fig. 9). We identified SNP rs12252 in our sequencing and compared the allele and genotype frequencies to allele and genotype frequencies from 1000 Genomes sequencing data from different populations (Supplementary Table 3). In addition, we used the most recent release of phased 1000 Genomes data³⁰ to impute the region surrounding SNP rs12252 to determine allele frequencies in the publicly available genotype data set of WTCCC1 controls ($n = 2,938$) and four previously published data sets genotyped from the Netherlands ($n = 8,892$) and Germany ($n = 6,253$)²². In the imputation, samples genotyped with Illumina 550k, 610k and 670k platforms were imputed against the June 2011 release of 1000 Genomes phased haplotypes using the Impute software³¹, version 2.1.2. Only individuals with European ethnicities (Europe (CEU), Finland (FIN), Great Britain (GBR), Spain/Iberia (IBS), Tuscany (TSI)) were included from the 1000 Genomes reference panel. Recommended settings were used for imputing the region 200 kb in either direction from the variants of interest, along with 1 Mb buffer size. The statistical significance of the allele frequencies was determined by Fisher's exact test.

We assessed for population stratification by principal component analysis. Genotype data from the WTCCC1 1958 Birth Cohort data set were obtained from the European Genotype Archive with permission, reformatted and merged with genotype data from the GenISIS study to match 113,819 SNPs present in both cohorts. Suspected strand mismatches were removed by identifying SNPs with more than 2 genotypes and using the LD method as implemented in Plink (v1.07)³², resulting in 105,362 matched SNPs. Quality control was applied in GenABEL version 1.6-9 to genotype data for these SNPs for the GenISIS cases and 1,499 individuals from WTCCC. Thresholds for quality control (deviation from Hardy-Weinberg equilibrium ($P < 0.05$), minor allele frequency (MAF) < 0.0005, call rate < 98% in all samples) were applied iteratively to identify all markers and subjects passing all quality control criteria, followed by principal component analysis using GenABEL. We tested for positive selection using both a haplotype-based test (|XP-EHH-max|) and allele frequency spectrum-based test statistics, namely the CLR²³⁻²⁵ on 10-kb windows across the entire genome as described previously^{30,33}. The three statistics were combined and the combined P value was plotted corresponding to the 10-kb windows.

26. de Wit, E. *et al.* Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. *Virus Res.* **103**, 155–161 (2004).
27. Hutchinson, E. C., Curran, M. D., Read, E. K., Gog, J. R. & Digard, P. Mutational analysis of *cis*-acting RNA signals in segment 7 of influenza A virus. *J. Virol.* **82**, 11869–11879 (2008).
28. Britten, K. M., Howarth, P. H. & Roche, W. R. Immunohistochemistry on resin sections: a comparison of resin embedding techniques for small mucosal biopsies. *Biotech. Histochem.* **68**, 271–280 (1993).
29. Ellis, J. *et al.* Virological analysis of fatal influenza cases in the United Kingdom during the early wave of influenza in winter 2010/11. *Eurosurveillance* **16**, 2–7 (2011).
30. The 1000 Genomes project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
31. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
32. Purcell, S. *et al.* PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
33. MacArthur, D. G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).

Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes

Wen Tang^{1,2,3}, Ram Kannan^{1,2,3}, Marco Blanchette^{2,4} & Peter Baumann^{1,2,3}

In most eukaryotes, the progressive loss of chromosome-terminal DNA sequences is counteracted by the enzyme telomerase, a reverse transcriptase that uses part of an RNA subunit as template to synthesize telomeric repeats. Many cancer cells express high telomerase activity, and mutations in telomerase subunits are associated with degenerative syndromes including dyskeratosis congenita and aplastic anaemia. The therapeutic value of altering telomerase activity thus provides ample impetus to study the biogenesis and regulation of this enzyme in human cells and model systems. We have previously identified a precursor of the fission yeast telomerase RNA subunit (TER1)¹ and demonstrated that the mature 3'-end is generated by the spliceosome in a single cleavage reaction akin to the first step of splicing². Directly upstream and partly overlapping with the spliceosomal cleavage site is a putative binding site for Sm proteins. Sm and like-Sm (LSm) proteins belong to an ancient family of RNA-binding proteins represented in all three domains of life³. Members of this family form ring complexes on specific sets of target RNAs and have critical roles in their biogenesis, function and turnover. Here we demonstrate that the canonical Sm ring and the Lsm2–8 complex sequentially associate with fission yeast TER1. The Sm ring binds to the TER1 precursor, stimulates spliceosomal cleavage and promotes the hypermethylation of the 5'-cap by Tgs1. Sm proteins are then replaced by the Lsm2–8 complex, which promotes the association with the catalytic subunit and protects the mature 3'-end of TER1 from exonucleolytic degradation. Our findings define the sequence of events that occur during telomerase biogenesis and characterize roles for Sm and Lsm complexes as well as for the methylase Tgs1.

In eukaryotes, seven Sm proteins (SmB, SmD1, 2 and 3, SmE, SmF and SmG) form a heteroheptameric complex at U-rich Sm-binding sites (AU₄₋₆GR) of various small nuclear RNAs (snRNAs) including the spliceosomal snRNAs U1, U2, U4 and U5 (refs 4, 5). Assembly of Sm proteins *in vivo* requires the help of motor neuron protein (SMN), mutations in which result in spinal muscular atrophy⁶. At least two Sm-like complexes have been characterized. The Lsm1–7 complex functions in messenger RNA (mRNA) degradation^{7,8} and the Lsm2–8 complex is involved in the maturation of various polymerase III transcripts^{9–11} and ribosomal RNAs¹². Purified Lsm2–8 binds to the 3'-terminal U-tract on U6, but not to the internal U-rich Sm sites in U1, U2, U4 and U5 snRNAs, illustrating that Sm and Lsm complexes have different sets of target RNAs⁹.

Sm-binding sites are also found near the 3'-ends of telomerase RNA subunits from diverse yeasts^{1,13–16} and are important for RNA processing and/or stability^{1,2,15}. Actual binding of Sm proteins has been demonstrated for TLC1, the telomerase RNA from *Saccharomyces cerevisiae*¹⁵, but the functional consequences of this interaction have remained largely unexplored. The Sm-binding site in TLC1 is located several nucleotides upstream of the mature 3'-end¹³. In contrast, spliceosomal cleavage of *Schizosaccharomyces pombe* TER1 truncates the putative Sm-binding site by one nucleotide², which may compromise stable association of the Sm ring with mature TER1. We therefore set

out to examine which proteins bind to the 3'-end of mature TER1, and to determine the function of the putative Sm site for TER1 biogenesis and stability.

A strategy was developed to examine the 3'-end of TER1 by massively parallel sequencing to obtain a quantitative measure of 3'-end sequence distribution and to identify the most abundant terminal sequences (Fig. 1a). This analysis revealed that, after spliceosomal cleavage, over 60% of TER1 molecules lost additional nucleotides at the 3'-end and terminate in a stretch of three to six uridines (Fig. 1b). The 3'-end of most of TER1 therefore resembles the 3'-end of U6 snRNA, which is bound by the Lsm2–8 complex. To test whether Sm or Lsm proteins associate with TER1, carboxy-terminal c-Myc epitope tags were inserted at the genomic loci of all Sm and Lsm proteins.

Immunoprecipitations were performed with a subset of strains that did not show overt growth defects, expressed tagged proteins and maintained telomeres (Supplementary Fig. 1). The snRNA U1 control specifically co-precipitated with Sm proteins, confirming that the epitope tags did not interfere with immunoprecipitation of RNP complexes (Fig. 1c). TER1 co-immunoprecipitated with all four Sm

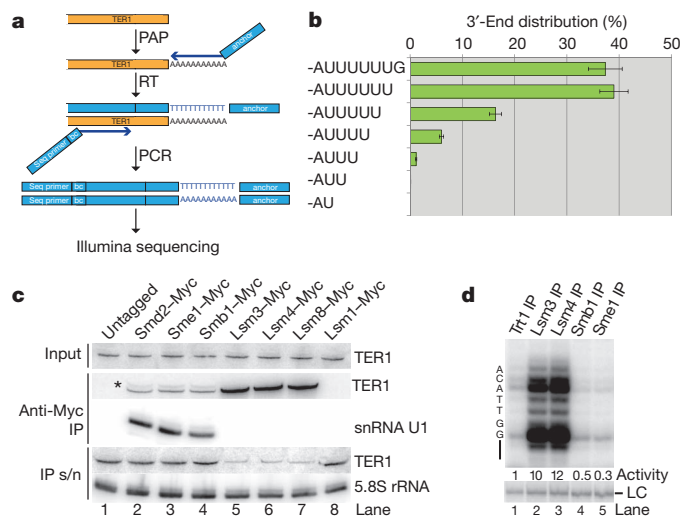


Figure 1 | TER1 RNA associates with Sm and Lsm proteins. **a**, Method used to map the 3'-end distribution of TER1 post spliceosomal cleavage. RNA is depicted in orange, DNA in blue. PAP, poly(A) polymerase; RT, reverse transcription; PCR, polymerase chain reaction; bc, barcode. **b**, Distribution of 3'-end positions in mature TER1 from wild-type cells. The average of four experiments is shown; error bars, standard deviation; a total of 23×10^6 sequences were scored. **c**, Northern blot of RNA isolated from immunoprecipitations with anti-c-Myc antibodies. Input and immunoprecipitation (IP) supernatant (s/n) represent 10% of the sample. An asterisk marks the position of the TER1 precursor. The lower band corresponds to the mature form of TER1. **d**, Telomerase activity assay performed on beads after c-Myc immunoprecipitation of tagged proteins as indicated above each lane. Activity was quantified relative to the Trt1 immunoprecipitate sample. A 100-mer [³²P]oligonucleotide was used as recovery and loading control (LC).

¹Howard Hughes Medical Institute, Kansas City, Missouri 64110, USA. ²Stowers Institute for Medical Research, Kansas City, Missouri 64110, USA. ³Department of Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, Kansas 66160, USA. ⁴Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas 66160, USA.

proteins tested (Fig. 1c, lanes 2–4, and Supplementary Fig. 2a), including members of each of the three Sm subcomplexes⁴. Strikingly though, several-fold more TER1 was recovered from Lsm immunoprecipitates resulting in an approximately 80% depletion of TER1 from the immunoprecipitation supernatant (Fig. 1c, lanes 5–7). TER1 precipitated with all subunits of the Lsm2–8 complex (Fig. 1c and Supplementary Fig. 2b), but not with Lsm1 (Fig. 1c, lane 8), the subunit specific to the Lsm1–7 complex.

To determine whether Sm and/or Lsm are associated with active telomerase, direct *in vitro* activity assays were performed on immunoprecipitates. Telomerase activity was detected in all samples, but was 20-fold higher in Lsm3 and 4 than Smb1 and Sme1 immunoprecipitates (Fig. 1d and Supplementary Fig. 2c). In part this can be explained by the lower recovery of telomerase with Sm proteins, as judged by quantification of telomerase RNA on northern blots (Supplementary Fig. 2c, d). However, even after normalization to the amount of TER1 in each immunoprecipitate, Lsm-associated telomerase activity was still 2.8-fold higher than that associated with Sm proteins. The simplest explanation for this observation is that a fraction of Sm-associated TER1 is not yet associated with the catalytic subunit of telomerase. Indeed, further experiments confirmed that Sm binding precedes Trt1 binding to TER1.

To gain insights into the functions of Sm and Lsm binding to telomerase we initially focused on the Sm association. For most characterized snRNAs, sequences downstream of the Sm-binding site are critical for Sm loading¹⁷. As the mature form of TER1 lacks such sequences, we tested whether the Sm complex was loaded onto the TER1 precursor before spliceosomal cleavage. Reverse transcription PCR (RT-PCR) confirmed that the precursor is indeed specifically associated with the Sm complex, but is undetectable in Lsm immunoprecipitations (Fig. 2a).

As the spliceosome contains Sm complexes, the TER1–Sm interaction may reflect binding of the spliceosome to the TER1 precursor. To test whether Sm proteins bind TER1 directly, we generated constructs with either a mutant 5′-splice-site or a deletion of the intron. Both mutant RNAs co-immunoprecipitated with Smb1 (Fig. 2b). In contrast, replacing the Sm-binding sequence with a random sequence (*ter1-sm6* mutant) reduced Sm association by 22-fold (Fig. 2c). Similarly, Lsm association was undetectable for *ter1-sm6* (Supplementary Fig. 3a). We therefore surmised that Sm and Lsm proteins directly bind to the previously identified site in TER1.

We next examined the effect of Sm binding on 3′-end processing by the spliceosome. Loss of Sm binding in the *ter1-sm6* mutant resulted in a sevenfold reduction in the processed form (Fig. 2d). Furthermore, a series of deletion mutants within the Sm site caused progressive inhibition of TER1 cleavage (Supplementary Fig. 3b), but not TER1 splicing (Supplementary Fig. 3c). Finally, introducing an eight-nucleotide spacer between the Sm site and 5′-cleavage-site also impaired processing (Fig. 2e). In summary, weakening or abolishing Sm association with the TER1 precursor reduces spliceosomal cleavage, indicating that Sm proteins promote 3′-end processing of TER1.

A conserved feature among yeast and mammalian telomerase RNAs is the post-transcriptional hypermethylation of the 5′-cap into a 2,2,7-trimethyl guanosine (TMG) form^{1,15,18}. Sm proteins were first implicated in promoting cap hypermethylation on U2 snRNA in *Xenopus* extract¹⁹. It was later shown *in vitro* that TMG-capping of human U1 requires the presence of Smb/B′-SmD3 (refs 4, 20). A screen for physical association with Sm proteins led to the identification of the methylase Tgs1 in budding yeast²¹. To elucidate the roles of Sm and/or Lsm in the hypermethylation of the 5′-cap on TER1, we tested which, if any, of these proteins interact with *S. pombe* Tgs1 (ref. 22) by two-hybrid analysis. Smd proteins scored positive, with Smd2 displaying the strongest interaction, and the other Sm proteins and all Lsm proteins showing weak or no interaction (Supplementary Fig. 4a). We next examined whether preventing Sm binding to TER1 affects cap hypermethylation. Whereas wild-type TER1 was readily precipitated

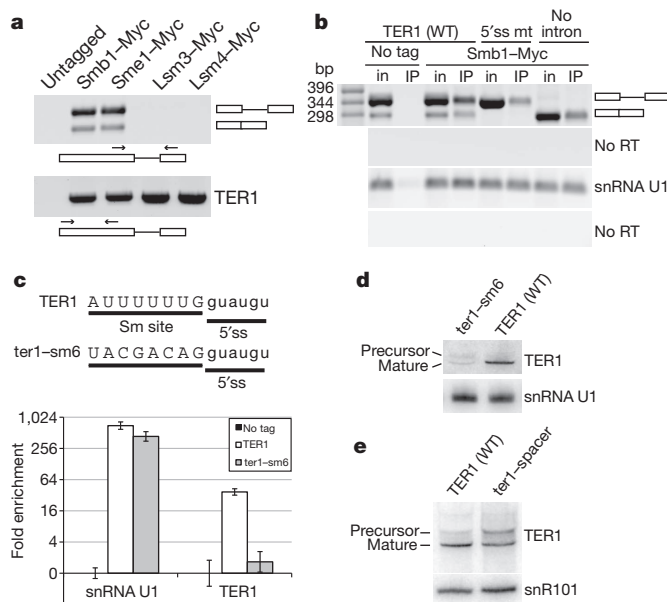


Figure 2 | Sm proteins associate with TER1 precursor and promote spliceosomal cleavage. **a**, RNA from anti-c-Myc immunoprecipitates was analysed by RT-PCR using primers in the first and second exon (primers represented by arrows in the schematic below the gel) to amplify the precursor form (upper panel). The primer pair also amplifies the spliced form (lower band in Sm immunoprecipitates). A primer pair in the first exon was used to visualize all forms of TER1 combined (lower panel). **b**, Sm association does not require spliceosome assembly on TER1. RT-PCR was performed on RNA purified from input (in) and anti-c-Myc immunoprecipitate beads (IP). Primers amplifying snRNA U1 were used as a positive control. **c**, The Sm-binding site (upper case) and 5′-splice-site (5′ss, lower case) for wild-type TER1 and the *ter1-sm6* mutant (MT). Replacing the Sm-binding site on TER1 (*ter1-sm6* mutant) compromises Sm association. RNA recovered from anti-c-Myc immunoprecipitates from untagged control and Smb1-Myc strains was quantified by real-time PCR. Data are plotted as enrichment over the untagged control. Error bars, standard error of triplicate experiments. **d**, Sm site mutation affects TER1 spliceosomal cleavage. Total RNA samples were analysed by northern blot for TER1 and snRNA U1. **e**, Increasing the distance between the Sm site and 5′-splice-site in the *ter1-spacer* mutant (AU₆GgccaugGU) impairs TER1 processing. Northern blot for TER1 and snRNA snR101 as loading control.

with a monoclonal antibody against the TMG cap, *ter1-sm6* recovery was at least 25-fold reduced (Fig. 3a and Supplementary Fig. 4b). Only the cleaved form of TER1 was recovered in TMG immunoprecipitations from wild-type cells, suggesting that spliceosomal cleavage precedes hypermethylation (Supplementary Fig. 4c). TER1 was not TMG-capped in a *tgs1Δ* strain, confirming that Tgs1 is the enzyme responsible for TER1 cap hypermethylation (Supplementary Fig. 4d).

In light of the reported increase in telomerase RNA in *tgs1Δ* budding yeast²³, we were surprised to observe a fivefold reduction in mature TER1 RNA in *tgs1Δ* compared with wild type in *S. pombe* (Fig. 3b). In addition, an increase in the precursor indicated a 3′-end processing defect. The viability of *tgs1Δ* cells ruled out a major splicing defect, but we consistently noted a small reduction in spliceosomal snRNAs isolated from *tgs1Δ* cells (Fig. 3b and data not shown). To differentiate between a processing defect and a direct effect of the TMG cap on TER1 stability, we mutated the spliceosomal cleavage site and inserted a hammerhead ribozyme sequence to generate the mutant *ter1-5′ssmut-HH* (Supplementary Fig. 4e). In this construct, processing of TER1 occurs independently of the spliceosome by ribozyme cleavage. When comparing *ter1-5′ssmut-HH* levels between wild-type and *tgs1Δ* cells, a twofold reduction was observed (Fig. 3b). Taken together, these results show that *tgs1Δ* affects TER1 processing by the spliceosome as well as TER1 stability. Consistent with the exquisite dosage sensitivity for telomerase RNA in diverse species^{24,25}, this reduction in TER1

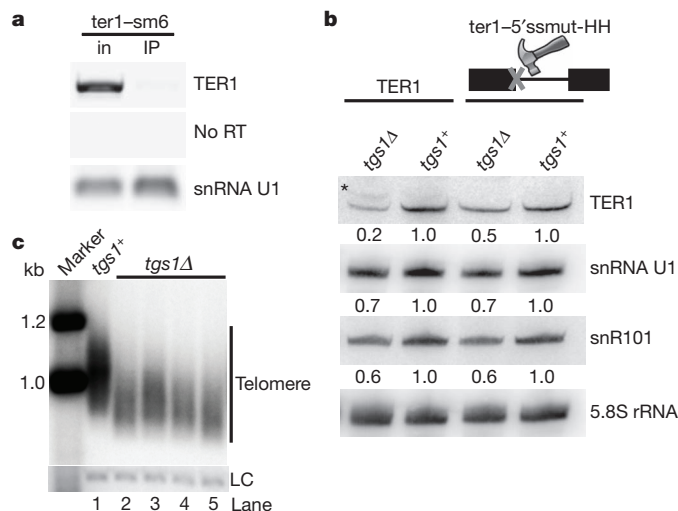


Figure 3 | Tgs1 modifies TER1 and is required for normal telomere maintenance. **a**, Loss of Sm site compromises TMG cap formation. RT-PCR amplifying all forms of TER1 and *ter1-sm6* mutant from anti-TMG immunoprecipitation (IP) and input (in) samples; snRNA U1 served as control. **b**, Bypass of spliceosomal cleavage reveals functions of Tgs1 in TER1 processing and stability. Northern blot analysis of TER1, snRNA U1, snR101 and 5.8S rRNA from total RNA prepared from wild-type and *tgs1Δ* strains harbouring either TER1 or the *ter1-5'ssmut-HH* mutant. An asterisk marks the position of the TER1 precursor. **c**, Deletion of *tgs1*⁺ causes telomere shortening. Telomere length was analysed by Southern blotting of EcoRI-digested genomic DNA from four independent *tgs1Δ* isolates and an otherwise isogenic *tgs1*⁺ strain. A probe for the *rad16* gene was used as a loading control (LC).

resulted in shorter telomeres (Fig. 3c). Neither telomerase activity nor Lsm association was reduced beyond the effects expected from the reduced steady-state level of TER1 (Supplementary Fig. 4f, g).

Most TER1 post-spliceosomal cleavage was bound by Lsm2–8, but a small fraction was associated with Sm proteins (Fig. 1c). To investigate whether this was indicative of a switch from Sm to Lsm binding, we examined the distribution of 3'-ends in each immunoprecipitation by massively parallel sequencing. Around 70% of Sm-bound TER1 post-cleavage terminated precisely at the spliceosomal cleavage site (Fig. 4a and Supplementary Fig. 5a). Enrichment of this form in the Sm-bound fraction is consistent with Sm proteins binding the TER1 precursor and remaining associated with TER1 until after cleavage and cap hypermethylation have occurred. In contrast, Lsm-associated TER1 predominantly terminated in U₃₋₆, indicating that a switch between Sm and Lsm binding occurs after spliceosomal cleavage and is associated with exonucleolytic processing (Fig. 4a and Supplementary Fig. 5b).

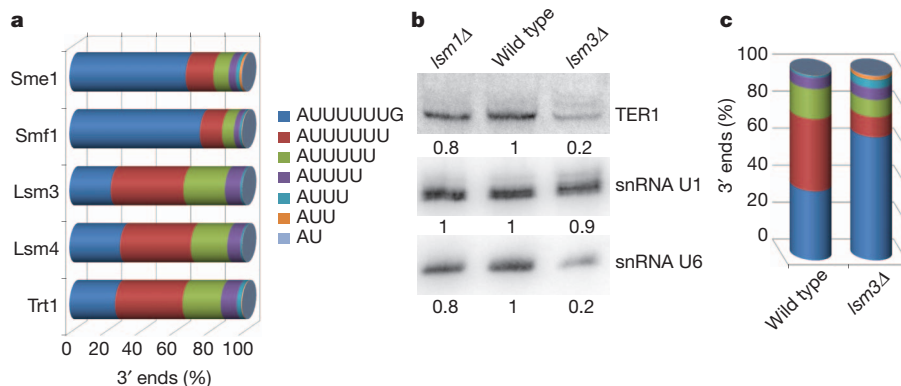


Figure 4 | Lsm proteins replace Sm and protect the 3'-end of TER1. **a**, 3'-End sequence distribution of TER1 from immunoprecipitation samples. **b**, Northern blot analysis from total RNA prepared from wild-type, *lsm1Δ* and *lsm3Δ* strains, quantified relative to wild type for each RNA. **c**, Specific loss of

Consistent with most telomerase activity being associated with Lsm2–8, the TER1 3'-end distribution from Trt1 immunoprecipitates was indistinguishable from that of Lsm-bound TER1.

The observation that loss of Sm binding coincided with the loss of terminal nucleotides led us to speculate that Lsm2–8 may function in protecting the 3'-end of TER1 against further exonucleolytic degradation. To test this, we attempted to generate Lsm deletion strains. Whereas most Lsm proteins are essential, *lsm1Δ* and *lsm3Δ* cells were viable. Consistent with a protective function for Lsm2–8, the levels of TER1 and U6 snRNA were reduced approximately fivefold in *lsm3Δ* cells (Fig. 4b). No such effect was seen when deleting *lsm1*, nor was the level of U1 snRNA reduced in *lsm3Δ* cells. The 3'-end sequence distribution for TER1 from total RNA of *lsm3Δ* cells closely resembled the Sm-bound fraction in wild type, whereas the Lsm-bound fraction was selectively lost in the mutant (Fig. 4c and Supplementary Fig. 5c). The viability of *lsm3Δ* cells further allowed us to confirm that cap hypermethylation is unaffected by the absence of Lsm consistent with Tgs1 acting on TER1 before Lsm binding (Supplementary Fig. 5d).

To verify independently a role for Lsm proteins in stabilizing TER1, we took advantage of the observation that Lsm binding requires a stretch of consecutive uridines⁹. In contrast, Sm binding tolerates other nucleotides in certain positions of the binding motif, as exemplified by the Sm-binding site in human U1 snRNA (AAUUUGUG). When the TER1 Sm site was mutated to reduce the number of consecutive uridines, the level of mature TER1 was decreased (Fig. 5a). We next precipitated Smb1, Lsm4 and Trt1 from wild type and strains containing the *ter1-SmU1* mutant. As expected, the mutation had little effect on the binding of Sm proteins (Fig. 5b). In fact, when normalized for the lower level of *ter1-SmU1* compared with wild type, recovery of *ter1-SmU1* with Smb1 was increased 1.6-fold. In contrast, Lsm binding was diminished by more than 20-fold. Most surprisingly, the interaction between the catalytic subunit Trt1 and telomerase RNA was also compromised in the *ter1-SmU1* mutant (Fig. 5b). The normalized recovery of *ter1-SmU1* with Trt1 was 15-fold lower than wild type, indicating that Lsm binding facilitates Trt1–TER1 association, possibly by inducing a conformational change in the RNA analogous to how binding of the p65 protein facilitates telomerase assembly in *Tetrahymena*^{26–28}. Consistent with the poor recovery of *ter1-SmU1* in Trt1 immunoprecipitations, *in vitro* telomerase activity was below the level of detection (Fig. 5c).

Analysis of the 3'-end sequence distribution for *ter1-SmU1* from total RNA revealed that most of the mutant RNA ends at the cleavage site (Supplementary Fig. 6). This form constituted close to 90% of *ter1-SmU1* in Smb1 immunoprecipitates. In contrast, Lsm4 and Trt1 immunoprecipitates predominantly recovered RNA ending in -AUUU and -AUUUG (Supplementary Fig. 6). These results further support that Trt1 preferentially associates with Lsm-bound telomerase RNA.

Lsm2–8-bound fraction of TER1 in *lsm3Δ* cells based on 3'-end sequence analysis from total RNA samples. The wild-type sample from Fig. 1b is included for comparison.

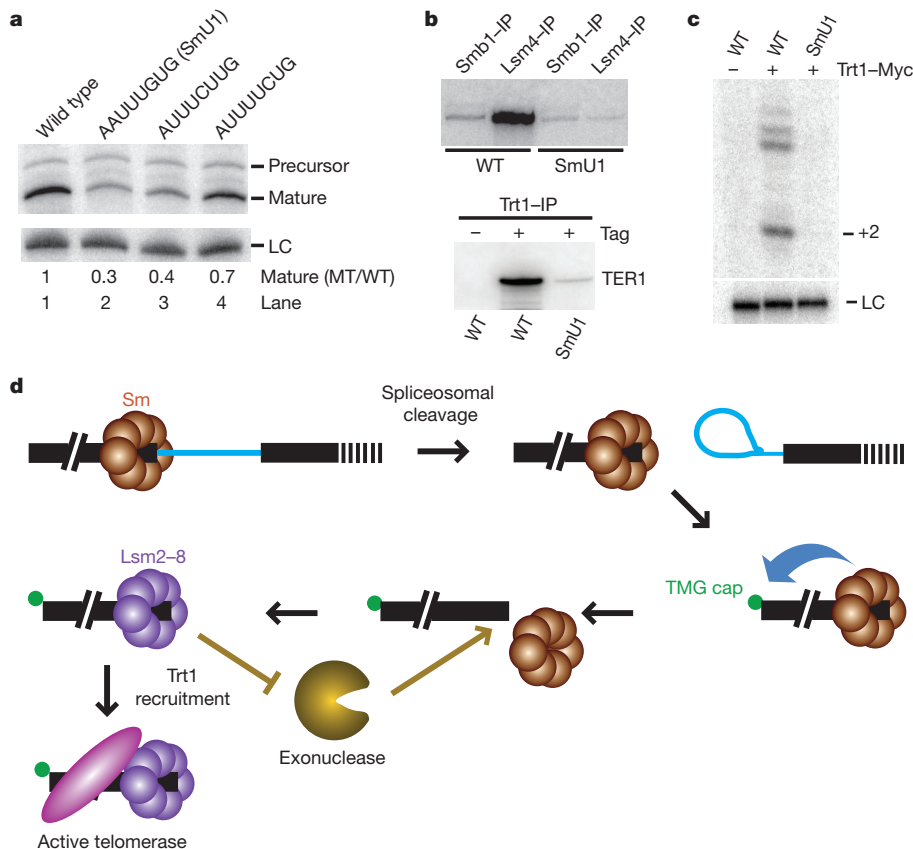


Figure 5 | Lsm binding to TER1 promotes telomerase assembly and protects TER1 from degradation. **a**, Northern blot for TER1. The indicated ratios of mutant (MT) to wild type (WT) are normalized to the loading control snR101 (LC). **b**, Northern blot for TER1 and the *ter1*-SmU1 mutant using RNA isolated from anti-c-Myc immunoprecipitations performed on extract from

They also confirm the role of Lsm in protecting the 3'-end of TER1 from further degradation, as diminished Lsm binding coincides with an overall reduction in telomerase RNA and a shift towards the form that is bound by Sm.

Taken together, our observations demonstrate that distinct populations of TER1 molecules associate with the Sm and Lsm complexes and suggest a sequence of events for TER1 biogenesis (Fig. 5d). The polyadenylated TER1 precursor is bound by the Sm complex, which promotes spliceosomal cleavage and subsequent 5'-cap hypermethylation by recruiting Tgs1. The Sm ring is then replaced by the Lsm2-8 complex, which protects TER1 from exonucleolytic degradation and promotes binding of the catalytic subunit.

Despite their structural similarity and related binding motifs, Sm and Lsm complexes have different modes of RNA binding and were thought to have distinct and non-overlapping sets of target RNAs. The finding that the TER1 precursor is exclusively associated with the Sm complex, whereas most mature TER1 is bound by Lsm2-8, revealed that biogenesis of telomerase RNA involves both Sm and Lsm complexes. Considering the central roles that Sm and Lsm proteins play in RNA metabolism, it will be important to determine whether biogenesis of other non-coding RNAs also involves Sm- and Lsm2-8-bound stages. Furthermore, it is interesting to note that several human Sm/Lsm proteins have been reported to co-purify with telomerase^{29,30}, raising the possibility that these proteins also function in TMG cap formation and telomerase assembly in metazoans.

METHODS SUMMARY

Myc epitope tags were integrated at the genomic loci and immunoprecipitations were performed in whole-cell extracts with anti-c-Myc antibodies. The different forms of telomerase RNA were detected by northern blotting and RT-PCR. The

strains harbouring Smb1-Myc, Lsm4-Myc or Trt1-Myc as indicated.

c, Telomerase activity assay performed on Trt1 immunoprecipitates from strains harbouring either wild type or *ter1*-SmU1. An untagged Trt1 strain was used as negative control. **d**, Sequence of events that occur during telomerase biogenesis.

distribution of 3'-ends was assessed at single nucleotide resolution by preparing libraries of oligo(A)-tailed telomerase RNA and massively parallel sequencing.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 September 2011; accepted 7 February 2012.

Published online 25 March 2012.

- Leonardi, J., Box, J. A., Bunch, J. T. & Baumann, P. TER1, the RNA subunit of fission yeast telomerase. *Nature Struct. Mol. Biol.* **15**, 26–33 (2008).
- Box, J. A., Bunch, J. T., Tang, W. & Baumann, P. Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature* **456**, 910–914 (2008).
- Wilusz, C. J. & Wilusz, J. Eukaryotic Lsm proteins: lessons from bacteria. *Nature Struct. Mol. Biol.* **12**, 1031–1036 (2005).
- Raker, V. A., Plessel, G. & Lührmann, R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *EMBO J.* **15**, 2256–2269 (1996).
- Patel, S. B. & Bellini, M. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* **36**, 6482–6493 (2008).
- Terns, M. P. & Terns, R. M. Macromolecular complexes: SMN — the master assembler. *Curr. Biol.* **11**, R862–R864 (2001).
- Tharun, S. *et al.* Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**, 515–518 (2000).
- Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M. & Seraphin, B. A Sm-like protein complex that participates in mRNA degradation. *EMBO J.* **19**, 1661–1671 (2000).
- Achsel, T. *et al.* A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation *in vitro*. *EMBO J.* **18**, 5789–5802 (1999).
- Mayes, A. E., Verdone, L., Legrain, P. & Beggs, J. D. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.* **18**, 4321–4331 (1999).
- Kufel, J., Allmang, C., Verdone, L., Beggs, J. D. & Tollervey, D. Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol. Cell. Biol.* **22**, 5248–5256 (2002).
- Kufel, J., Allmang, C., Petfalski, E., Beggs, J. & Tollervey, D. Lsm proteins are required for normal processing and stability of ribosomal RNAs. *J. Biol. Chem.* **278**, 2147–2156 (2003).

13. Dandjinou, A. T. *et al.* A phylogenetically based secondary structure for the yeast telomerase RNA. *Curr. Biol.* **14**, 1148–1158 (2004).
14. Gunisova, S. *et al.* Identification and comparative analysis of telomerase RNAs from *Candida* species reveal conservation of functional elements. *RNA* **15**, 546–559 (2009).
15. Seto, A. G., Zaug, A. J., Sobel, S. G., Wolin, S. L. & Cech, T. R. *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature* **401**, 177–180 (1999).
16. Webb, C. J. & Zakian, V. A. Identification and characterization of the *Schizosaccharomyces pombe* TER1 telomerase RNA. *Nature Struct. Mol. Biol.* **15**, 34–42 (2008).
17. Yong, J., Kasim, M., Bachorik, J. L., Wan, L. & Dreyfuss, G. Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. *Mol. Cell* **38**, 551–562 (2010).
18. Jady, B. E., Bertrand, E. & Kiss, T. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.* **164**, 647–652 (2004).
19. Mattaj, J. W. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell* **46**, 905–911 (1986).
20. Plessel, G., Fischer, U. & Lührmann, R. m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) *in vitro*: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol. Cell Biol.* **14**, 4160–4172 (1994).
21. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonne, R. Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell* **9**, 891–901 (2002).
22. Hausmann, S., Ramirez, A., Schneider, S., Schwer, B. & Shuman, S. Biochemical and genetic analysis of RNA cap guanine-N2 methyltransferases from *Giardia lamblia* and *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **35**, 1411–1420 (2007).
23. Franke, J., Gehlen, J. & Ehrenhofer-Murray, A. E. Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J. Cell Sci.* **121**, 3553–3560 (2008).
24. Harrington, L. Making the most of a little: dosage effects in eukaryotic telomere length maintenance. *Chromosome Res.* **13**, 493–504 (2005).
25. Mozdy, A. D. & Cech, T. R. Low abundance of telomerase in yeast: implications for telomerase haploinsufficiency. *RNA* **12**, 1721–1737 (2006).
26. Berman, A. J., Gooding, A. R. & Cech, T. R. Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol. Cell Biol.* **30**, 4965–4976 (2010).
27. O'Connor, C. M. & Collins, K. A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol. Cell Biol.* **26**, 2029–2036 (2006).
28. Stone, M. D. *et al.* Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* **446**, 458–461 (2007).
29. Fu, D. & Collins, K. Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev.* **20**, 531–536 (2006).
30. Fu, D. & Collins, K. Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol. Cell* **28**, 773–785 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Shuman for the *tgs1Δ* strain, J. A. Box, J. T. Bunch, S. Hartnett and R. M. Helston for technical assistance, the Molecular Biology Core Facility for site-directed mutagenesis and sequencing, D. P. Baumann and R. M. Helston for proofreading the manuscript, and all members of the Baumann laboratory for discussions. This work was funded in part by the Stowers Institute for Medical Research. R.K. is supported by an award from the American Heart Association, and P.B. is an Early Career Scientist with the Howard Hughes Medical Institute.

Author Contributions P.B. and W.T. conceived the study and designed the experiments; W.T. performed most of the experiments with some assistance from those acknowledged and P.B.; R.K. contributed to the characterization of Sm mutants and analysed telomere length of Myc-tagged strains. M.B. wrote the script for sequence data analysis and provided advice; W.T., R.K. and P.B. analysed the data, and W.T. and P.B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.B. (peb@stowers.org).

METHODS

Yeast strains and constructs. The genotypes of all strains used in this study are listed in Supplementary Table 1. Strains expressing c-Myc-tagged Sm and Lsm proteins were constructed in strain PP138 as described³¹. Mutants *ter1-sm6*, *ter1-smΔG*, *ter1-smΔUG*, *ter1-smΔU₂G*, *ter1-smΔU₃G* and *ter1-5'ssmut-HH* were integrated at the *ter1* genomic locus by gene replacement. Other *ter1* mutants were generated in the context of plasmid pJW10 using the QuikChange II XL site-directed mutagenesis kit (Stratagene) and introduced into PP407, PP694 or PP695 as described¹.

Yeast two-hybrid analysis. Yeast two-hybrid analysis used the Matchmaker GAL4 Two Hybrid System 3 (Clontech). Briefly, *tgsl*⁺ cDNA was cloned into the vector pGBKT7, and each full-length *lsm* and *sm* cDNA was cloned into pGADT7. Plasmids were co-transformed into the yeast strain AH109 and positive transformants were selected on SD–Leu–Trp plates. Interactions were analysed by plating threefold serial dilutions of overnight cultures onto SD–Leu–Trp–His–Ade plates. Plates were incubated for three days at 30 °C.

Telomere length analysis and telomerase activity assay. Cells were propagated for at least 80–100 generations and telomere length was analysed by Southern blotting as described³². Telomerase activity assays were performed on Sepharose beads as described^{1,33} after immunoprecipitation from cell extracts of strains harbouring Myc-tagged Trt1, Sm or Lsm proteins.

Immunoprecipitation and RNA isolation. *S. pombe* cells were grown in yeast extract supplements³² and 6 l of cell suspension were collected by centrifugation at a density of 5×10^6 cells per millilitre. Cells were washed in TMG(300) (10 mM Tris-HCl, pH 8.0, 1 mM magnesium chloride, 10% (v/v) glycerol, 300 mM sodium acetate), the pellet was resuspended in two packed cell volumes of TMG(300) plus supplements (5 µg ml⁻¹ chymostatin, 5 µg ml⁻¹ leupeptin, 1 µg ml⁻¹ pepstatin, 1 mM benzamidin, 1 mM DTT, 1 mM EDTA and 0.5 mM PMSF) and the suspension was frozen in liquid nitrogen. Cells were lysed under liquid nitrogen in a 6850 cryogenic mill (SPEX CertiPrep) with eight 2 min cycles at an impactor rate of 10 per second and a 2 min cooling time between cycles. The lysed cell powder was transferred into a 50 ml tube and allowed to thaw on ice for 30 min. Cell extracts were cleared by two rounds of centrifugation at 14,000g for 7 min and frozen in liquid nitrogen for storage at –80 °C. The concentration of proteins in the whole-cell extract was measured by Bradford protein assay. For c-Myc immunoprecipitation, monoclonal anti-c-Myc antibody (20 µg, Sigma) was incubated with 150 µl protein A/G agarose slurry (Calbiochem) in phosphate buffered saline at room temperature for 30 min. Beads were washed three times with TMG(300) plus supplements and whole-cell extract (1.2 ml) was added at a concentration of 5 mg ml⁻¹ together with RNasin (40 U, Promega), Tween 20 (0.1%) and heparin (1 mg ml⁻¹). For immunoprecipitation of TMG-capped RNAs, anti-TMG antibody (3 µg, Calbiochem) was bound to 50 µl protein A/G agarose slurry (Calbiochem), washed with TMG(300) and 150 µg total *S. pombe* RNA was added in 0.7 ml TMG(300). Samples were incubated on a rotator at 4 °C for 4 h, then washed three times with TMG(300) plus supplements and 0.1% Tween 20 and once with TMG(50) (as TMG(300) but only 50 mM sodium acetate). Protease inhibitors were omitted for TMG immunoprecipitations. RNA was isolated by treatment with proteinase K (2.0 mg ml⁻¹ in 0.5% (w/v) SDS, 40 mM EDTA, 20 mM Tris-HCl, pH 7.5) at 50 °C for 15 min, followed by extraction with acidic phenol and ethanol precipitation. RNA was then analysed by northern blotting, RT-PCR and 3'-end sequencing.

RNA analysis. RNA isolation and northern blotting were performed as described² except that Biotyne Nylon Transfer Membrane (Pall Corporation) was used and samples shown in Fig. 5a were treated with RNaseH in the presence of oligonucleotides BLoli1043 (AGGCAGAAGACTCACGTACACTGCAC), BLoli1275 and PBoli560 (GCGGAATTCT₁₈) to obtain better separation of precursor and mature form. The TER1 probe was generated as described²; other RNAs were detected using 5'-[³²P]DNA oligonucleotides as follows: GCTGCAGAACTCATGCCAGGTA AGT (snRNA U1), CGCTATTGTATGGGGCCTTTAGATTCTTA (snRNA snR101), CTTTCATCGATGCGAGAGCCAAGAGATCCGT (5.8S rRNA) and GCAGTGCATCCTTGTGCAGGGGCCA (snRNA U6).

Semi-quantitative RT-PCR was performed as described previously² with primers BLoli1275 (CGGAAACGGAATTCAGCATGT) and BLoli1020 (CAAA CAATAATGAACGTCCTG) amplifying the intron-spanning region, and PBoli918 (ACAACGGACGAGCTACACTC) and BLoli1006 (CATTTAAGTGC TTGTCTAGATCACAACG) amplifying a region in the first exon. BLoli2051 (GACCTTAGCCAGTCCACAGTTA) and BLoli2101 (ACCTGGCATGAGTTTC TGC) were used to amplify snRNA U1.

For quantitative real-time RT-PCR, reverse transcription for input and immunoprecipitated RNA were performed with antisense primer BLoli2860 (TGCTCAGACCAAGTGAAAAA) and BLoli2051. Real-time PCR was performed in triplicate 12.5 µl reactions using Power SYBR Green PCR Master Mix (Applied Biosystems) according to the manufacturer's instructions. BLoli2860 and BLoli2859 (GGATCAAAGCTTTTGCTTGT) were used to amplify the first exon of TER1. BLoli2051 and BLoli2101 were used to amplify snRNA U1. The qRT-PCR results were imported into Microsoft Excel and the average value and standard deviation of triplicate cycle threshold (C_t) values were calculated. Enrichment of immunoprecipitation is represented by ΔC_t (C_t value (immunoprecipitation sample) minus C_t value (input)) relative to the untagged control samples. Error bars in the graph represent the positive and negative range of the standard error of the mean.

3'-end cloning. DNase-treated total RNA samples (2.5 µg) or immunoprecipitated and purified RNA was incubated with poly(A) polymerase (600 U, US Biologicals), RNase inhibitor (RNasin, 40 U) and ATP (0.5 mM) in 20 µl reactions at 30 °C for 30 min. The reaction volume was increased to 35.5 µl by the addition of the oligonucleotide Boli2327 (CAAGCAGAAGACGGCATACGA(T)₁₈, 125 pmol) and dNTP mix (25 nmol), and reactions were incubated at 65 °C for 3 min followed by slow cooling to room temperature. The reaction volume was then adjusted to 50 µl with first strand buffer (Invitrogen), dithiothreitol (5 mM), RNasin (40 U) and Superscript III reverse transcriptase (200 U, Invitrogen), and reactions were incubated at 50 °C for 60 min. RNaseH (5 U, NEB) was added and incubation was continued at 37 °C for 20 min. Aliquots (3 µl) of this reaction were used in PCR with Taq polymerase (5 U, NEB), primers (GTTCAGAGTTCTACAGTCCGAC GATC#GCAAAATGTTAAAGGAACG) and Boli2330 (CAAGCAGAAGAC GGCATACGA) (200 nM each, ## represents a two-nucleotide barcode used for multiplexing) under the following conditions: 3 min at 94 °C followed by 10 cycles of 30 s at 94 °C, 45 s at 55 °C and 60 s at 72 °C, followed by 7 min at 72 °C. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and eluted with 46 µl elution buffer. In the second round of PCR, 23 µl of the eluted product was amplified with BLoli2329 (AATGATACGGCGACCGACAGG TTCAGAGTTCTACAGTCCGA) and BLoli2330 (200 nM each) under the following conditions: 3 min at 94 °C followed by 29 cycles of 30 s at 94 °C, 45 s at 55 °C and 60 s at 72 °C, followed by 7 min at 72 °C. PCR products were separated by electrophoresis on 1.5% agarose gels, and bands of the correct size were excised and purified. The concentration of the PCR products was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies) and further adjusted to 10 nM for massively parallel sequencing using Illumina sequencing technology. Reads were analysed using a custom script written in BioPerl to filter for those that contained the TER1 sequence (GCAAAAN₁₀AACG) and to sort the reads into different bins based on the two-nucleotide barcodes. The nucleotide sequence between GCAAAAN₁₀AACG and the oligo(A) sequence resulting from the poly(A) polymerase treatment represents the end of TER1 and was used to determine the 3'-end sequence distribution at single nucleotide resolution. Further analysis and graphs were prepared in Microsoft Excel.

- Bahler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**, 943–951 (1998).
- Bunch, J. T., Bae, N. S., Leonardi, J. & Baumann, P. Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. *Mol. Cell. Biol.* **25**, 5567–5578 (2005).
- Haering, C. H., Nakamura, T. M., Baumann, P. & Cech, T. R. Analysis of telomerase catalytic subunit mutants in vivo and in vitro in *Schizosaccharomyces pombe*. *Proc. Natl Acad. Sci. USA* **97**, 6367–6372 (2000).

DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation

Pierre Close^{1,2}, Philip East³, A. Barbara Dirac-Svejstrup¹, Holger Hartmann⁴, Mark Heron⁴, Sarah Maslen⁵, Alain Chariot², Johannes Söding⁴, Mark Skehel⁵ & Jesper Q. Svejstrup¹

Alternative messenger RNA splicing is the main reason that vast mammalian proteomic complexity can be achieved with a limited number of genes. Splicing is physically and functionally coupled to transcription, and is greatly affected by the rate of transcript elongation^{1–3}. As the nascent pre-mRNA emerges from transcribing RNA polymerase II (RNAPII), it is assembled into a messenger ribonucleoprotein (mRNP) particle; this is the functional form of the nascent pre-mRNA and determines the fate of the mature transcript⁴. However, factors that connect the transcribing polymerase with the mRNP particle and help to integrate transcript elongation with mRNA splicing remain unclear. Here we characterize the human interactome of chromatin-associated mRNP particles. This led us to identify deleted in breast cancer 1 (DBC1) and ZNF326 (which we call ZNF-protein interacting with nuclear mRNPs and DBC1 (ZIRD)) as subunits of a novel protein complex—named DBIRD—that binds directly to RNAPII. DBIRD regulates alternative splicing of a large set of exons embedded in (A + T)-rich DNA, and is present at the affected exons. RNA-interference-mediated DBIRD depletion results in region-specific decreases in transcript elongation, particularly across areas encompassing affected exons. Together, these data indicate that the DBIRD complex acts at the interface between mRNP particles and RNAPII, integrating transcript elongation with the regulation of alternative splicing.

The composition of mRNP particles has been the subject of a number of studies, using a variety of approaches (for example, ref. 5). There are likely to be different types of mRNP particles with distinct compositions and interaction partners. We sought to purify native mRNP particles and interacting proteins from the chromatin in which they are generated and in which they are active in co-transcriptional processes. As a starting point, we generated HEK293 cells expressing near-normal levels of Flag-tagged heterogeneous nuclear ribonucleoprotein (hnRNP) A1, an abundant hnRNP protein in human cells⁶. hnRNP A1 shuttles between the nucleus and the cytoplasm⁷, but at steady state it is mainly nuclear and concentrated in chromatin (Fig. 1a), from where it can be released by RNase A treatment (Fig. 1b, compare lanes 2 and 4). We used DNase I digestion and mild sonication to release mRNP particles from chromatin for purification. RNase inhibitors were present during the whole process, outlined in Fig. 1c. mRNP particles isolated by this approach are predominantly of nuclear (chromatin) origin (Supplementary Fig. 1). Native mRNP particles and their interacting partners were purified from chromatin isolated from ~10⁸ nuclei (Fig. 1d). Only two major bands (namely the added, proteinaceous RNase inhibitors) were detected after purification from control cells, whereas numerous proteins were detected in hnRNP A1-Flag elutions (Fig. 1d). These represent a heterogeneous mixture of core mRNP particle subunits and proteins interacting with such particles. Individual protein bands were excised and identified by mass spectrometric analysis (Fig. 1d; see also

Supplementary Fig. 2a). Most of the known ‘core’ mRNP proteins, such as the hnRNP proteins, were present in the purified fraction, confirming the biological relevance of this approach. Many other pre-mRNA processing proteins were also identified, including splicing factors, ATP-dependent RNA helicases and a substantial number of mRNA 3′-end

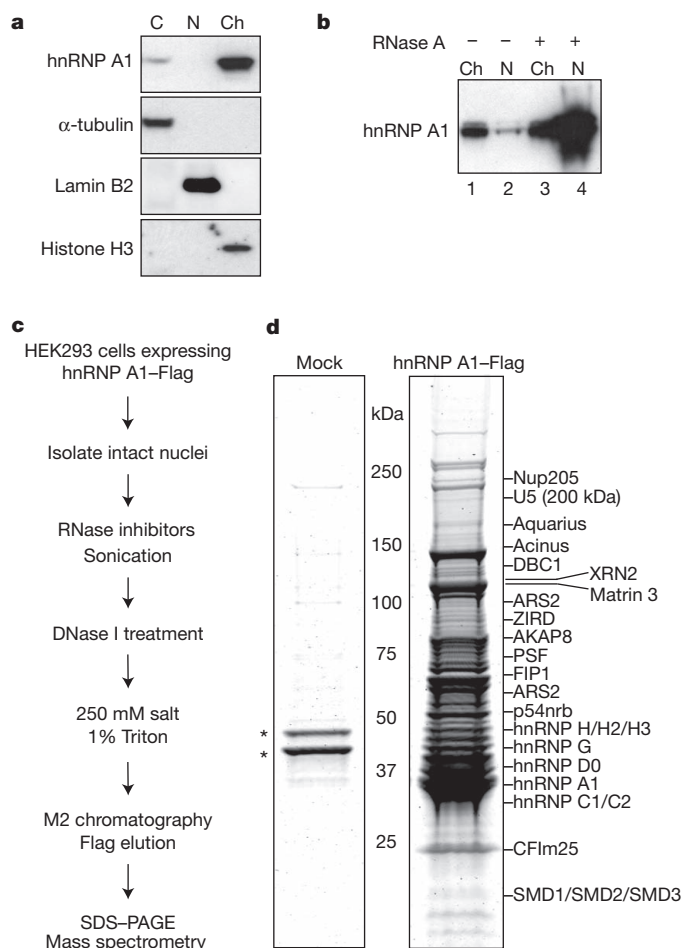


Figure 1 | Purification of nascent nuclear mRNP particles. **a**, Western blot analysis of cytoplasm (C), nucleoplasm (N) and chromatin (Ch), with α -tubulin, lamin B2 and histone H3 as controls for different fractions. **b**, Fractionation as in part **a**, but RNase A was added to the nuclear lysis buffer in the fractions indicated. **c**, Outline of the purification procedure. **d**, Equal fractions of the M2 chromatography eluates from control (Mock) and hnRNP A1-Flag separated by 4–12% SDS-PAGE and stained with SYPRO ruby. Asterisks mark RNase inhibitor proteins. Some of the identified proteins are indicated on the right.

¹Mechanisms of Transcription Laboratory, Cancer Research UK London Research Institute, Clare Hall Laboratories, South Mimms EN6 3LD, UK. ²Unit of Medical Chemistry, GIGA-Signal Transduction, GIGA-R, University of Liege, CHU, Sart-Tilman, 4000 Liege, Belgium. ³Bioinformatics & Biostatistics Group, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK. ⁴Gene Center and Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany. ⁵Protein Analysis and Proteomics Laboratory, Clare Hall Laboratories, Cancer Research UK, London Research Institute, South Mimms EN6 3LD, UK.

processing and termination factors. Co-immunoprecipitation experiments confirmed the RNA-dependent interaction of some of these proteins with hnRNP A1-Flag (Supplementary Fig. 2b).

We next focused on two proteins that had not previously been connected to mRNP particles or mRNA processing. One of these, DBC1, is best known for its association with, and regulation of, the sirtuin-like deacetylase SIRT1 (refs 8, 9). We also investigated the uncharacterized zinc-finger-containing protein ZNF326. Stable cell lines expressing near-normal levels of Flag-tagged versions of these proteins were established, and co-immunoprecipitation experiments confirmed that both DBC1 and ZNF326 interact with mRNP particles in an RNA-dependent manner (Supplementary Fig. 3a–f). Furthermore, we discovered that ZNF326 and DBC1 associate directly, in an RNA-independent manner (Fig. 2a, e). For this reason, hereafter we refer to ZNF326 as ZNF-protein interacting with nuclear mRNPs and DBC1 (ZIRD).

We previously identified DBC1 as an RNAPII-interacting protein in another proteomic screen¹⁰, making it a particularly interesting candidate. Co-immunoprecipitation experiments confirmed that RNAPII associates with DBC1-Flag in an RNA-independent manner (Fig. 2b). Furthermore, ZIRD was detected in RNAPII (RPB3-Flag) purifications, and this interaction was also RNA-independent (Fig. 2c). In further support of a ZIRD–RNAPII interaction, ZIRD-Flag also co-immunoprecipitated RNAPII (Fig. 2d). In contrast, we failed to detect an interaction between hnRNP A1 and RNAPII under the same conditions (Fig. 2c, middle panel, and data not shown), although co-immunoprecipitation experiments after formaldehyde crosslinking indicated that, as expected, the proteins are in close proximity *in vivo* (Supplementary Fig. 4). Together, these results indicate that DBC1 and ZIRD are not part of the core mRNP particle, but that they might work at the interface between the mRNP particle and RNAPII.

Others reported that DBC1 interacts with SIRT1 (refs 8, 9). Although we observed that DBC1 co-precipitated SIRT1, endogenous ZIRD and ZIRD-Flag did not (Fig. 2e, and data not shown). SIRT1 is also absent from hnRNP-A1-containing mRNP particles (Supplementary Fig. 3g). This indicates that ZIRD and DBC1 form a complex that lacks SIRT1. To characterize the ZIRD–DBC1 interaction further, ZIRD-Flag was purified. Size-exclusion chromatography of highly purified material showed that ZIRD-Flag and DBC1 are part of a salt-stable ~800-kDa complex (Fig. 2f) that also co-purified on MonoQ (data not shown). As expected, SIRT1 is not part of this protein complex (Fig. 2g, and data not shown). We named it the DBC1–ZIRD complex (DBIRD).

DBC1 and ZIRD interact with RNAPII in crude extracts (Fig. 2b–d). To investigate whether this interaction is direct, the DBIRD complex was characterized by gel filtration after mixing with an excess of RNAPII. In the absence of RNAPII, the DBIRD complex peaked in fractions 13–15 (Fig. 2f, upper two panels). However, when mixed with RNAPII (Fig. 2f, lower three panels), DBIRD complex elution shifted to earlier fractions, peaking in fraction 10 with a sub-fraction of RNAPII, whereas polymerase alone peaked in fractions 17–19 (~500 kDa), as expected. The DBIRD complex thus seems to form a bridging complex that interacts with both mRNP particles and RNAPII. Interestingly, DBIRD also interacted with mRNP particles lacking hnRNP A1 (Fig. 2h), pointing to a general bridging role.

To examine the role of the DBIRD complex in transcription-associated processes *in vivo*, we analysed the transcriptome of cells that had been depleted for DBC1 or ZIRD by RNA interference (RNAi) (Supplementary Fig. 5a). Total mRNA was hybridized to GeneChip Human Exon 1.0 ST arrays, on which the abundance of individual exons can be analysed independently. In the absence of ZIRD, a greater than 1.5-fold increase in exon inclusion was observed

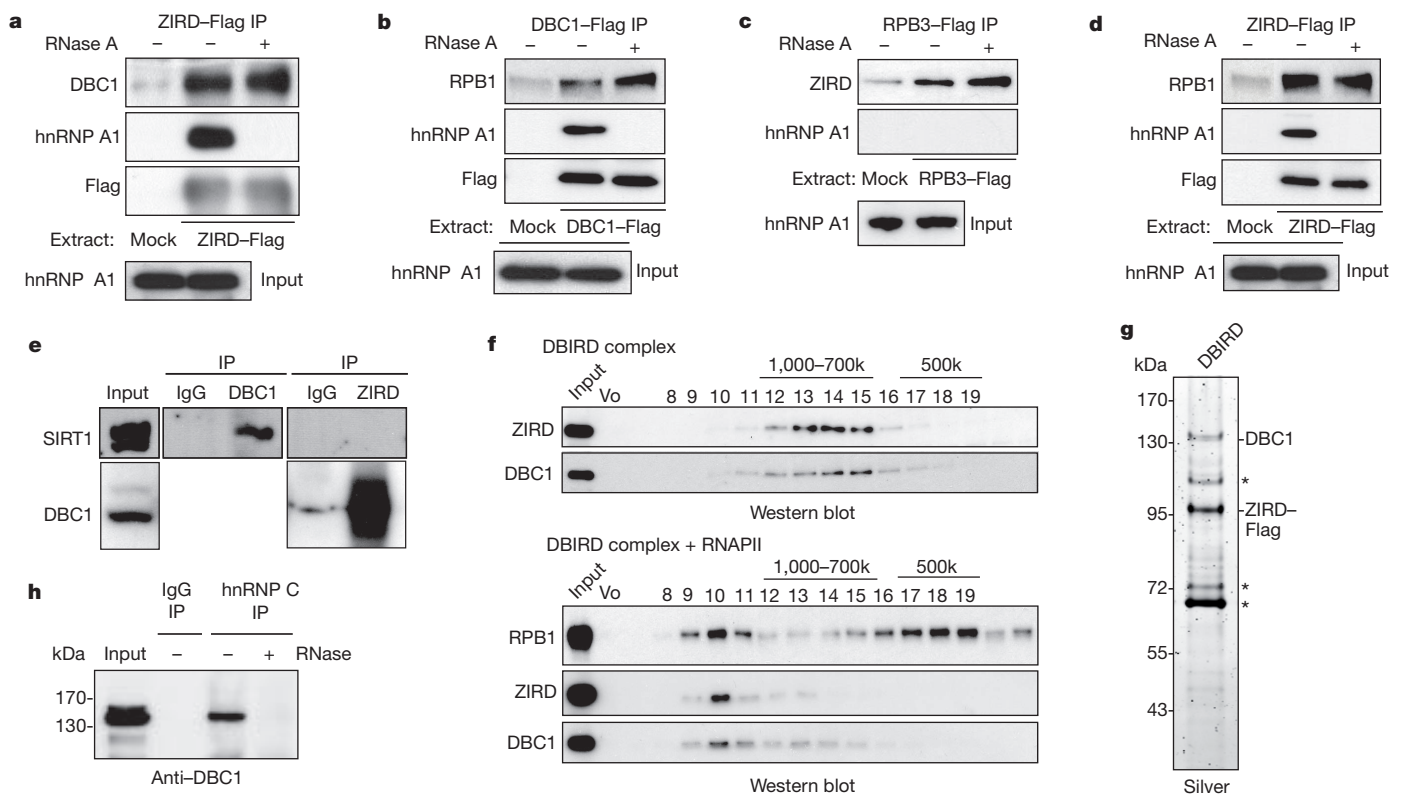


Figure 2 | DBC1 and ZIRD form a stable complex that binds RNAPII. **a**, Western blot analysis of anti-Flag immunoprecipitates (IPs) from ZIRD-Flag cells. **b**, As in **a**, but for DBC1-Flag cells. **c**, As in **a**, but RPB3-Flag cells. **d**, As in **a**, but detecting RPB1. **e**, Western blot analysis of anti-DBC1 and anti-ZIRD immunoprecipitates. IgG, immunoglobulin-γ. **f**, DBIRD analysed

by size-exclusion chromatography with (lower three panels) or without (upper two panels) RNAPII in approximately fivefold molar excess. Vo, void volume fraction. **g**, Silver stain of DBIRD. Asterisks indicate DBC1 and ZIRD degradation products. **h**, Western blot analysis of hnRNP-C-containing mRNP particles from mouse CB3 cells lacking hnRNP A1 (ref. 20).

in more than 2,800 situations, whereas exon exclusion was observed in only 390 cases (Supplementary Table 1a). The absence of DBC1 led to increased inclusion of an exon in 796 cases (Supplementary Table 1b) and, notably, most of these events were also on the list of ZIRD-dependent exon inclusions (567 out of 796 (71%); P -value for shared exons = 6.705×10^{-261} ; Supplementary Table 1c), which strongly supports a close functional relationship between the two factors and provides confidence in the genome-wide alternative splicing data sets. The effect was at the level of alternative splicing, as depletion of ZIRD or DBC1 only affected the expression of a very small number of genes (Supplementary Fig. 6).

A full list of inclusion events observed in both DBC1- and ZIRD-depleted cells is in Supplementary Table 1c. Sample results were confirmed by quantitative PCR with reverse transcription (RT-PCR) (Supplementary Fig. 7). To investigate whether DBIRD was present at affected exons, we performed RNA immunoprecipitation experiments¹¹. DBC1 and ZIRD bound the relevant exon in mRNAs from seven tested genes, whereas other regions (or control transfer RNA) were not detected or detected to a much lower extent (Fig. 3a, b and Supplementary Fig. 8). Interestingly, some exons of the β -actin gene (whose splicing was unaffected by DBIRD depletion) had considerable levels of DBIRD complex as well (Supplementary Fig. 8), indicating that the interaction of DBIRD with mRNA is not invariably correlated with DBIRD-dependent splicing changes.

To investigate the mechanism underlying exon inclusion, we first searched for sequence motifs in the DNA encompassing the included exons, but failed to uncover motifs other than those known to typify splice junctions. We then looked for nucleotide patterns that might be over-represented in the sequences surrounding the included exons by counting how often each of the 1,024 possible 5-base oligonucleotides occurred. Interestingly, (A + T)-rich 5-base oligonucleotides were markedly enriched around included exons (Fig. 3c). The frequencies of the four nucleotides in the regions around the splice sites were also analysed. A and T were strongly over-represented around the splice sites of DBIRD-affected exons, as well as across the exons themselves (Fig. 3d). The observed difference in A + T content is sufficient to explain the over-representation of (A + T)-rich 5-base oligonucleotides (Supplementary Fig. 9).

The (A + T)-rich DNA surrounding the affected exons might influence fundamental aspects of transcription. Indeed, A- and

T-tracts are difficult for RNAPII to transcribe, as they constitute very efficient elongation pause sites *in vitro*^{12,13}. To investigate the effect of DBIRD on transcript elongation, we performed RNAPII chromatin-immunoprecipitation (ChIP) analysis after DBIRD knockdown. For a control, we also knocked down SIRT1 (Supplementary Fig. 5b). Remarkably, although overall transcription of *RAD50* and *SLC36A4* is not affected (see Supplementary Fig. 7), depletion of DBC1 or ZIRD (but not SIRT1) markedly affected RNAPII transcription distinctively in regions encompassing affected exons (Fig. 4 and Supplementary Fig. 11). Quantification of newly produced mRNA by bromo-UTP incorporation supported the idea that elongation rates were decreased in these regions (Supplementary Fig. 10). DBIRD depletion also affected RNAPII density at other genes whose splicing was exon-specifically affected, whereas little or no change in RNAPII density was observed at the unaffected β -actin control gene, even at exons that had an elevated DBIRD level (Supplementary Figs 11 and 12; compare to Supplementary Fig. 8).

Our data support the idea that the DBIRD complex functions at the interface between core mRNP particles and RNAPII, affecting local transcript elongation rates and alternative splicing at a subset of (A + T)-rich exon-intron junctions (Supplementary Fig. 13). Notably, several studies have shown that the rate of RNAPII elongation affects the efficiency of splicing, with slow elongation favouring exon inclusion¹⁻³. Therefore, one possible explanation for our data is that the DBIRD complex acts as an elongation factor that facilitates transcript elongation across (A + T)-rich regions, and thereby affects alternative splicing of exons in these regions. It has also been suggested that exons in the nascent pre-mRNA become tethered to the elongating transcription complex^{14,15}. Given that DBIRD binds both mRNPs and RNAPII, it might affect such tethering, and thereby affect splicing.

DBC1 has been implicated in tumorigenesis as a potential tumour suppressor, regulating apoptosis and cell survival¹⁶. Whether the role of DBC1 in the DBIRD complex and alternative splicing affects tumorigenesis is an interesting possibility, particularly in light of the recent finding that genes encoding components of the splicing machinery are often mutated in myelodysplastic syndromes and related disorders¹⁷. ZIRD has not previously been characterized in human cells, but its mouse homologue, ZAN75, is highly expressed in neuronal tissues¹⁸, suggesting that regulation of the DBIRD complex might contribute to

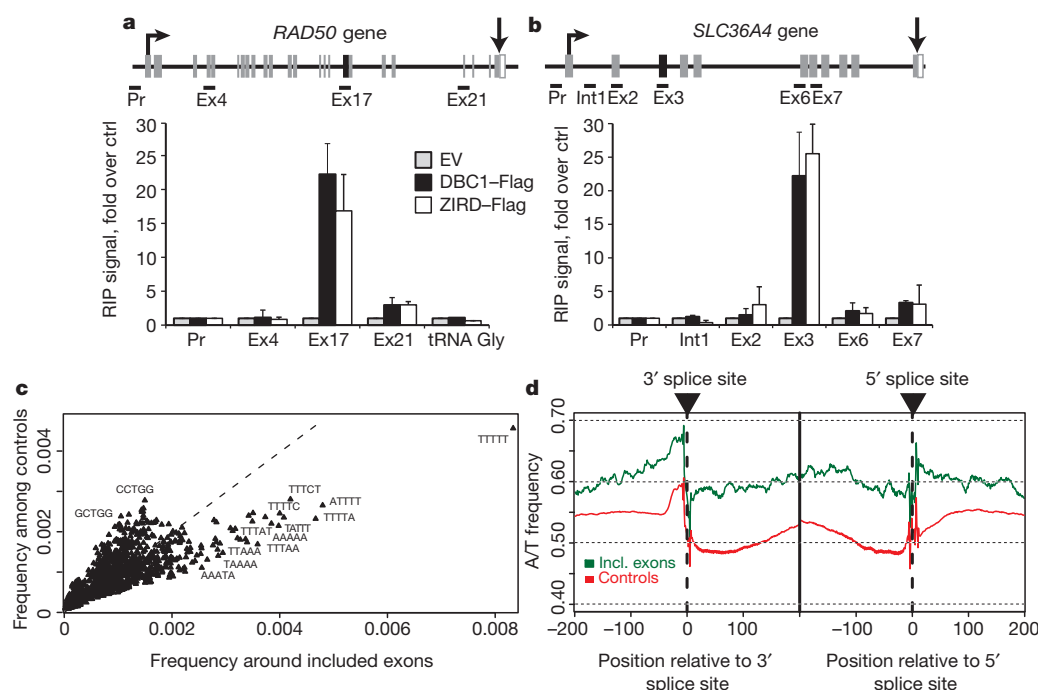


Figure 3 | DBIRD affects alternative splicing and is present at the affected exons. **a, b**, RNA immunoprecipitation (RIP) from crosslinked control (EV, ctrl), DBC1-Flag or ZIRD-Flag cells, analysed by quantitative PCR (qPCR). Control reactions lacking reverse transcriptase were always included (not shown). Error bars indicate standard deviations according to the Poisson statistic; $n = 3$. EV, empty vector; Ex, exon; Int, intron; Pr, promoter. Arrows indicate the start of transcription and the end of the coding region, respectively. **c**, Frequency of 5-base oligonucleotides in the regions around splice sites of affected (x axis) versus unaffected (y axis) exons. Diagonal line marks equal frequencies in the positive and negative set. **d**, Frequency of A or T upstream and downstream from splice sites of included exons (green) and unaffected control exons (red). The solid black line marks the meeting point of upstream and downstream data sets.

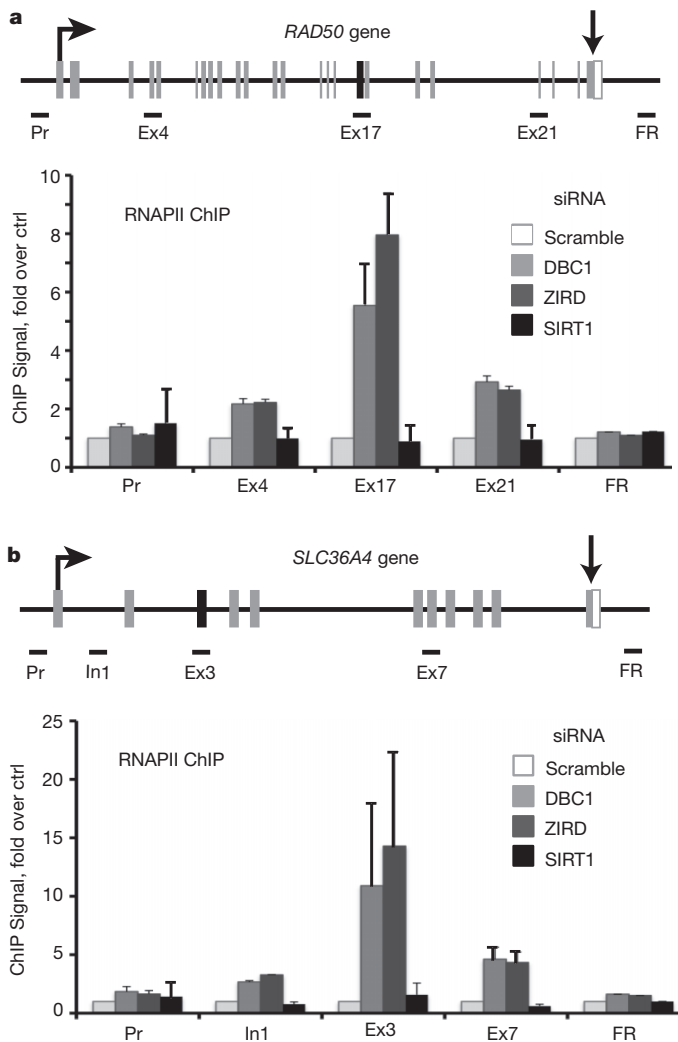


Figure 4 | DBC1 and ZIRD link exon skipping to RNAPII transcription. **a**, *RAD50* gene and qPCR primers (upper panel). RNAPII ChIP using cells transfected with control (scramble), *DBC1*, *ZIRD* or *SIRT1* Stealth siRNAs (lower panel). ChIP signals were normalized with inputs. Signals in control cells were set to 1 at each position, and values obtained from factor-depleted cells were expressed relative to these signals. Errors bars denote standard deviation; $n = 3$. **b**, Same as in **a**, but for *SLC36A4*. Supplementary Fig. 12 shows the same data in a format in which gene positional information is maintained. FR, flanking region. Arrows indicate the start of transcription and the end of the coding region, respectively.

tissue-specific splicing. Other proteins with homology to ZIRD and DBC1 exist in the human genome, raising the possibility that other DBIRD-like complexes are specific for other sets of genes or exons, or are involved in other transcription-related nuclear events.

METHODS SUMMARY

Open reading frames encoding hnRNP A1, DBC1 and ZIRD were cloned into pIRESpuro (Clontech) with a carboxy-terminal Flag tag. HEK293 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS in 5% CO₂ at 37 °C. For proteomic analysis, nuclei were isolated from hnRNP A1-Flag cells. These were sonicated and treated with DNase I, and the sample was cleared by centrifugation and the supernatant was subjected to M2 agarose (Sigma) chromatography. Proteins were eluted with 3× Flag peptide and mass spectrometry was performed as has been described elsewhere¹⁰. DBIRD was purified by M2 agarose chromatography from nuclease-treated nuclear extract from cells expressing ZIRD-Flag. DBIRD was analysed by MonoQ, or size exclusion chromatography with or without an excess of RNAPII. Stealth short interfering RNAs (siRNAs) were double transfected in HEK293 cells using lipofectamine 2000

(Invitrogen). For microarray analysis, RNA was hybridized on Human Exon 1.0 ST arrays (Affymetrix) using standard techniques (bioinformatics analysis described in Methods). For assessment of exon abundance and transcript expression, quantitative RT-PCR was performed using primers against affected and unaffected exons. Primer details are given in Supplementary Table 2. RNA immunoprecipitation and ChIP assays were performed as described^{11,19}.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 May 2011; accepted 7 February 2012.

Published online 25 March 2012.

- Kornblihtt, A. R. Coupling transcription and alternative splicing. *Adv. Exp. Med. Biol.* **623**, 175–189 (2007).
- Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R. & Misteli, T. Epigenetics in alternative pre-mRNA splicing. *Cell* **144**, 16–26 (2011).
- Perales, R. & Bentley, D. "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* **36**, 178–191 (2009).
- Dreyfuss, G., Kim, V. N. & Kataoka, N. Messenger-RNA-binding proteins and the messages they carry. *Nature Rev. Mol. Cell Biol.* **3**, 195–205 (2002).
- Chen, Y. I. et al. Proteomic analysis of *in vivo*-assembled pre-mRNA splicing complexes expands the catalog of participating factors. *Nucleic Acids Res.* **35**, 3928–3944 (2007).
- Kiledjian, M., Burd, C. G., Görlach, M., Portman, D. S. & Dreyfuss, G. in *Frontiers in Molecular Biology* (eds Mattaj, I. & Nagai, K.) 127–149 (Oxford Univ. Press, 1994).
- Piñol-Roma, S. & Dreyfuss, G. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**, 730–732 (1992).
- Kim, J. E., Chen, J. & Lou, Z. DBC1 is a negative regulator of SIRT1. *Nature* **451**, 583–586 (2008).
- Zhao, W. et al. Negative regulation of the deacetylase SIRT1 by DBC1. *Nature* **451**, 587–590 (2008).
- Aygün, O., Svejstrup, J. & Liu, Y. A RECQ5–RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc. Natl Acad. Sci. USA* **105**, 8580–8584 (2008).
- Selth, L. A., Close, P. & Svejstrup, J. Q. Studying RNA-protein interactions *in vivo* by RNA immunoprecipitation. *Methods Mol. Biol.* **791**, 253–264 (2011).
- Sigurdsson, S., Dirac-Svejstrup, A. B. & Svejstrup, J. Q. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol. Cell* **38**, 202–210 (2010).
- Saeki, H. & Svejstrup, J. Q. Stability, flexibility, and dynamic interactions of colliding RNA polymerase II elongation complexes. *Mol. Cell* **35**, 191–205 (2009).
- Dye, M. J., Gromak, N. & Proudfoot, N. J. Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849–859 (2006).
- Fong, N., Ohman, M. & Bentley, D. L. Fast ribozyme cleavage releases transcripts from RNA polymerase II and aborts co-transcriptional pre-mRNA processing. *Nature Struct. Mol. Biol.* **16**, 916–922 (2009).
- Kim, J. E., Chen, J. & Lou, Z. p30 DBC is a potential regulator of tumorigenesis. *Cell Cycle* **8**, 2933–2936 (2009).
- Yoshida, K. et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
- Lee, J. Y. et al. Characterization of a zinc finger protein ZAN75: nuclear localization signal, transcriptional activator activity, and expression during neuronal differentiation of P19 cells. *DNA Cell Biol.* **19**, 227–234 (2000).
- Close, P. et al. Transcription impairment and cell migration defects in elongator-depleted cells: implication for familial dysautonomia. *Mol. Cell* **22**, 521–531 (2006).
- Ben-David, Y., Bani, M. R., Chabot, B., De Koven, A. & Bernstein, A. Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol. Cell Biol.* **12**, 4449–4455 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by grants from Cancer Research UK and the European Research council (ERC) (to J.Q.S.), by an European Molecular Biology Organization (EMBO) long-term fellowship and by the Fonds Leon Fredericq foundation (to P.C.). We thank the Molecular Biology Core Facility at the Paterson Institute (Manchester) and Cell Services at London Research Institute for analysis and help. We thank B. Chabot for CB3 cells, and members of the Svejstrup laboratory, P. Verrijzer and T. H. Jensen for comments on the manuscript. P.C. is a senior research assistant at the Belgian National Foundation for Scientific Research (FNRS).

Author Contributions P.C. and A.B.D.-S. performed experiments, and S.M. and M.S. did mass spectrometry analyses. P.E., H.H., M.H. and J.S. performed bioinformatic analyses. P.C. and J.Q.S. designed the study, analysed the experimental data and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Gene-expression and splicing data sets have been deposited in the GEO database under accession code GSE35480. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.Q.S. (j.svejstrup@cancer.org.uk).

METHODS

Plasmids and antibodies. Open reading frames encoding human hnRNP A1, DBC1 and ZIRD (ZNF326) were cloned into pIRESpuro (Clontech) with a Flag tag at the C terminus. Antibodies used were rabbit anti-Flag, mouse anti-Flag M2 and mouse anti-hnRNP C (Sigma); mouse anti-pCTD mAb 4H8 (Millipore); rabbit anti-lamin B2 (Acris); rabbit anti-ZNF326 and mouse anti-hnRNP A1 (Santa Cruz Biotechnology); and rabbit anti-DBC1 and rabbit anti-SIRT1 (Bethyl Laboratories).

Cell culture, stable-cell-line establishment and stealth siRNA transfection. HEK293 cells were grown in DMEM containing 10% FBS in 5% CO₂ at 37 °C. To generate HEK293 stably expressing a Flag-tagged protein, cells were transfected with the relevant pIRESpuro construct and selected in 1 µg ml⁻¹ puromycin (Sigma). Cells were maintained in selecting media for 3 weeks, and surviving cells were used for experiments after transgene expression was checked.

Stealth siRNAs were double transfected in HEK293 cells using lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Protein and RNA expression was checked 48 h after the second transfection. Stealth siRNA anti-ZIRD RNAi sequences were: 5'-CGGAGGUAGUUAUGGUGGUCGAUUU-3' (sense); 5'-AAAUCCGACCAUAACUACCUCCG-3' (antisense). Stealth siRNA anti-DBC1 RNAi sequences were: 5'-CCAUCUGUGACUCCUAGAACUCCA-3' (sense); 5'-UGGAGUUCUAGGAAGUCACAGAUGG-3' (antisense). For stealth siRNA anti-SIRT1 we used validated stealth siRNA (Invitrogen; oligo ID VHS0609). For control siRNA we used Stealth siRNA negative control med GC (Invitrogen; L2935-300).

Immunopurification of native mRNPs. 10⁸ cells stably expressing hnRNP A1-Flag were lysed with cytoplasmic lysis buffer (10 mM Tris HCl (pH 7.9), 340 mM sucrose, 3 mM CaCl₂, 2 mM Mg(OAc)₂, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 0.5% tertigol-type NP-40, protease inhibitors and 1 µl ml⁻¹ RNasin Ribonuclease inhibitor (Promega)), and intact nuclei were pelleted by centrifugation at 3,500g for 15 min. Nuclei were washed with cytoplasmic lysis buffer without NP-40 and then resuspended in DNase I buffer (20 mM HEPES (pH 7.9), 10% glycerol, 1.5 mM MgCl₂, 1 mM DTT, protease inhibitors and 1 µl ml⁻¹ RNasin). After ten strokes in a Dounce homogenizer, nuclei were sonicated using Bioruptor (Diagenode) before DNase I (Sigma) was added to the buffer and incubated for 30 min at room temperature. Buffer was then adjusted to a final concentration of 250 mM KOAc and 1% Triton X-100. The sample was cleared by centrifugation at 20,000g for 30 min and the supernatant was collected. For negative control purification, the same extracts were prepared from the same amount of untagged cells.

The sample was then applied to M2 agarose beads (Sigma) and incubated for 4 h at 4 °C. After binding, beads were washed extensively with washing buffer (20 mM HEPES (pH 7.9), 250 mM KOAc, 1% Triton X-100, 10% glycerol, 3 mM EDTA, 1 mM DTT, protease inhibitors and 1 µl ml⁻¹ RNasin Ribonuclease inhibitor)¹⁰. Finally, proteins were eluted by using Flag elution buffer (20 mM HEPES (pH 7.9), 100 mM KOAc, 3 mM EDTA, 1 mM DTT, 200 µg ml⁻¹ 3× Flag peptide, protease inhibitors and 1 µl ml⁻¹ RNasin Ribonuclease inhibitor). Eluates were resolved by 4–12% bis-Tris-gradient SDS-PAGE and revealed by SYPRO Ruby staining (Invitrogen).

Mass spectrometric analysis. Protein samples were reduced, alkylated and digested with trypsin, using the Janus liquid handling system (PerkinElmer). The digests were subsequently analysed by liquid chromatography tandem mass spectrometry on an LTQ Orbitrap XL mass spectrometer (ThermoScientific). The resulting data were searched against a protein database (UniProt KB) using the Mascot search engine programme (Matrix Science)²¹. All data were analysed manually.

Purification of the DBC1-ZIRD complex. Nuclei from approximately 10⁹ ZIRD-Flag cells were isolated, washed as above and then sonicated (Bioruptor, 30-s on-off cycles, max intensity for 15 min). Nucleic acids were digested by adding 10,000 units per ml Benzonase (Novagen) and 30 µg ml⁻¹ RNase A (Sigma), and incubating at 4 °C for 1 h. The nuclear extract was then adjusted to a final concentration of 250 mM KOAc and the insoluble fraction was removed by centrifugation (20,000g for 30 min). The supernatant was used for Flag-M2 chromatography. After extensive washes (60 column volumes of 20 mM HEPES-KOH (pH 7.9), 0.5% Triton X-100, 10% glycerol and 250 mM KOAc), bound proteins were eluted in the same buffer as above but containing 0.5 mg ml⁻¹ 3× Flag peptide (and 100 mM KOAc). The complex was then dialysed into buffer A (20 mM Tris-HCl (pH 7.9), 10% glycerol and 0.01% NP-40) containing 100 mM NaCl, before MonoQ PC 1.6/5 (GE Healthcare) or size exclusion (Superose 6 PC 3.2/30; GE Healthcare) chromatography. Proteins were eluted from MonoQ by a salt gradient from 0.1 to 1 M NaCl in buffer A. For size-exclusion chromatography, samples were loaded in buffer B (20 mM HEPES-KOH, 0.01% NP40, 10% glycerol and 250 mM KOAc) with or without pre-incubation on ice with a five- to tenfold molar excess of purified mammalian RNAPII, purified as described²². Fifty-mico-

litre fractions were collected, and sizes were estimated by running protein size markers (Biorad) in parallel.

Microarray analysis. RNAs were extracted using RNeasy Kit (Qiagen), DNase-I-treated on the column, labelled and hybridized on Human Exon 1.0 ST arrays (Affymetrix) using standard techniques. Three independent experiments were performed and used as real triplicate for data analysis. We processed core-probe-level signals using robust multiarray average (RMA) implemented in APT (apt-1.10.0, Affymetrix) to generate quantile-normalized probe-set and gene-level signal estimates. Probe set to transcript cluster metagrouping was obtained from Affymetrix. We removed control probe sets from further analysis. For the gene-level analysis, genes displaying a coefficient of variance of less than 0.05 were assumed to be uninformative and were removed. We determined transcriptional effects (DBC1 versus control, ZIRD versus control and DBC1 versus ZIRD) by linear model, moderating the *t*-statistics by empirical Bayes shrinkage. We selected differential genes using a 0.05 *P*-value threshold using a nested F method. The analysis was carried out using the limma package from Bioconductor version 2.3 (ref. 23).

To identify putative alternative splicing events we first filtered probe sets to reduce false positive events. We removed all probe sets that did not localize to unique loci in the genome (Affymetrix annotation). We removed all probe sets from transcripts identified as not expressed in a given condition, as it is not possible to determine alternative splicing events against an untranscribed background. We defined a transcript as not being expressed if less than half of its member probe sets had a detection *P* value (detection above background value (dabg)) of less than 0.05 across all replicates. We also removed probe sets displaying a dabg value of greater than 0.05 in all replicates in both the conditions being considered. Genes with less than three probe sets after filtering were also removed from the analysis. To identify pairwise alternative splicing events we fitted transcript-cluster-specific linear models to probe set signal estimates and tested for significant interactions between each probe set and gene-level signal estimates across pairwise conditions using a 0.01 *P*-value threshold and a nested F method. In cases in which multiple probe sets mapped to a single exon, only probe sets that had significant interactions were included in the results. Once again the analysis was carried out using the limma package from Bioconductor²³.

Computational analysis of sequences around affected exons. To investigate sequence features that might explain the differential enrichment of certain exons after knockdown of ZIRD or DBC1, we selected a set of 505 exons that were 1.5-fold enriched in both the ZIRD and the DBC1 knockdowns and which were not the first or the last exon (to avoid including 5' and 3' untranslated regions in the following analyses). As a negative control set, we selected a further 3,877 exons that were unaffected after depletion of ZIRD and that were not the first or the last exon in the gene. We then prepared a positive and a negative sequence set for both the 5' and 3' splice sites. For the 5' splice site, the positive set contained the regions from -200 nucleotides to +50 nucleotides around the 5' splice sites of the 505 enriched exons. The negative control set contained corresponding regions for the 3,877 unaffected exons. Similarly, the positive set for the 3' splice site contained the region from -50 nucleotides to +200 nucleotides around the 3' splice sites of the 505 enriched exons, and the negative set contained the corresponding regions of the 3,877 unaffected exons.

The *P* values in Supplementary Fig. 9 were obtained in the following way. Suppose *ab* to be a dinucleotide and *i* a position around a splice site. We calculated the frequency $P_{\text{exp}}(i, ab)$ of *ab* that we would expect given the frequency of *ab* in the negative set, $P_{\text{neg}}(i, ab)$, and given the frequencies in the positive (P_{pos}) and negative (P_{neg}) sets of mononucleotide *a* at position *i*, and *b* at position *i* + 1, respectively:

$$P_{\text{exp}}(i, ab) = P_{\text{neg}}(i, ab) \frac{P_{\text{pos}}(i, a)P_{\text{pos}}(i + 1, b)}{P_{\text{neg}}(i, a)P_{\text{neg}}(i + 1, b)}$$

The *P* values for the observed number of dinucleotides *ab* at position *i* were calculated by approximating the binomial distribution with a normal distribution:

$$P \text{ value} = \sum_{k=K}^N \binom{N}{k} p^k (1-p)^{N-k} = \frac{1}{2} \operatorname{erfc} \left(\frac{K - Np}{\sqrt{Np(1-p)}} \right)$$

where *N* is the number of sequences in the set of affected exons, *K* is the number of observed dinucleotides *ab* at position *i*, *Np* is the expected number of dinucleotides at *i*, $p = P_{\text{exp}}(i, ab)$, *k* is the summation value, and *erfc* is the complementary error function.

Quantitative RT-PCR. RNAs were extracted using RNeasy Kit (Qiagen) and DNase-I-treated on the column, and 1 µg of total RNA was retro-transcribed using random primers and a first-strand complementary DNA synthesis kit (Fermentas). Quantitative RT-PCR was performed using SYBR Green detection. Specific primers against the alternatively spliced exon and unaffected exons were designed to assess exon abundance and transcript expression. Sequences are available on request.

Chromatin immunoprecipitation. ChIP assays were performed as described¹⁹, using 4H8 antibody (Abcam) or IgG antibody as a negative control, and incubation for 1 h with protein G/Herring sperm DNA. The precipitated DNA fragments were analysed using real-time PCR with SYBR Green detection. Input DNA was analysed simultaneously and used for normalization.

RNA immunoprecipitation. RNA immunoprecipitation was performed essentially as described^{11,24}. Briefly, 10⁷ HEK293 cells were crosslinked in 1% PFA for 10 min, quenched with 0.125 M glycine for 5 min, washed twice in PBS and collected in lysis buffer (50 mM Tris-HCl (pH 8), 1% SDS, 10 mM EDTA, and 50 U per 500 µl RNasin and protease inhibitors). Samples were sonicated for 15 min at maximum power using Bioruptor (Diagenode) to obtain RNA fragments 200–600 bases long, cleared by centrifugation and diluted tenfold in dilution buffer (20 mM Tris-HCl (pH 8), 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 50 U ml⁻¹ RNasin and protease inhibitors). Samples were pre-cleared by incubation with 100 µl of IgG agarose beads (Sigma) for 2 h at 4 °C. Anti-Flag M2 Resin (Sigma) was added and samples were incubated for 4 h at 4 °C. Immunoprecipitates were washed three times in wash buffer 150 (20 mM Tris-HCl (pH 8), 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS, RNasin (50 U ml⁻¹) and protease inhibitors), once in wash buffer 500 (same as above, but with 500 mM NaCl), once in LiCl (10 mM Tris-HCl (pH 8), 250 mM LiCl, 0.5% NP-40, 0.1% Deoxycholate, 1 mM EDTA, 50 U ml⁻¹ RNasin and protease inhibitors) and once in TE100 buffer (TE containing 100 mM NaCl, 50 U ml⁻¹ RNasin, and protease inhibitors). Immunocomplexes were eluted three times in 150 µl elution buffer (TE containing 100 mM NaCl, 200 µg ml⁻¹ 3× Flag peptide, 50 U per ml RNasin) and incubated for 30 min at 37 °C with 1 µl of proteinase K. Crosslinking was then reversed by adding 9 µl of 5 M NaCl and incubating the samples at 65 °C for 1 h. Nucleic acids were extracted with phenol chloroform, then precipitated with ethanol and any remaining DNA was eliminated with Turbo DNase (Ambion) treatment. RNAs

were then reverse transcribed using random primers and the cDNAs were used for subsequent PCR reactions using relevant primers.

Nuclear run-on analysis. Cells were rinsed in PBS, then in buffer A (20 mM Tris-HCl (pH 7.4), 5 mM MgCl₂, 0.5 mM EGTA, 25% glycerol and 1 mM phenylmethylsulphonyl fluoride) and permeabilized in buffer A containing 0.02% Triton X-100 for 3 min at room temperature. The nascent RNA labelling reaction was carried out in buffer A containing 2 mM ATP, 0.5 mM GTP, 0.5 mM CTP, 0.2 mM BrUTP (Sigma) and 25 U ml⁻¹ RNasin (Promega) for 15 min at 37 °C. In control reactions, normal UTP was used instead of BrUTP. After BrU incorporation, cells were rinsed twice in PBS and total RNA from nuclei of both labelled and control samples was isolated using TriPure Isolation Reagent (Roche). BrdU antibody (2 µg) (Sigma; also recognizes BrU) was pre-incubated with 20 µl of Protein G magnetic beads (Invitrogen) per experimental condition. RNA was then heated at 80 °C for 10 min and incubated with the beads at room temperature for 1 h with gentle shaking. The beads were washed five times in PBS containing 0.1% polyvinylpyrrolidone and RNasin (20 U per 200 µl), the RNA bound to the beads was eluted and the contaminant DNA was eliminated with Turbo DNase (Ambion). RNA was then reverse transcribed using random primers and the cDNAs were used for subsequent PCR reactions using relevant primers.

21. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
22. Hu, X. *et al.* A Mediator-responsive form of metazoan RNA polymerase II. *Proc. Natl Acad. Sci. USA* **103**, 9506–9511 (2006).
23. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
24. Gilbert, C. & Svejstrup, J. Q. RNA immunoprecipitation for determining RNA-protein associations *in vivo*. *Curr. Protoc. Mol. Biol.* **75**, 27.4.1–27.4.11 (2006).

Differential positioning of adherens junctions is associated with initiation of epithelial folding

Yu-Chiun Wang^{1,2}, Zia Khan^{3,4†}, Matthias Kaschube^{4†} & Eric F. Wieschaus^{1,2}

During tissue morphogenesis, simple epithelial sheets undergo folding to form complex structures. The prevailing model underlying epithelial folding involves cell shape changes driven by myosin-dependent apical constriction¹. Here we describe an alternative mechanism that requires differential positioning of adherens junctions controlled by modulation of epithelial apical-basal polarity. Using live embryo imaging, we show that before the initiation of dorsal transverse folds during *Drosophila* gastrulation, adherens junctions shift basally in the initiating cells, but maintain their original subapical positioning in the neighbouring cells. Junctional positioning in the dorsal epithelium depends on the polarity proteins Bazooka and Par-1. In particular, the basal shift that occurs in the initiating cells is associated with a progressive decrease in Par-1 levels. We show that uniform reduction of the activity of Bazooka or Par-1 results in uniform apical or lateral positioning of junctions and in each case dorsal fold initiation is abolished. In addition, an increase in the Bazooka/Par-1 ratio causes formation of ectopic dorsal folds. The basal shift of junctions not only alters the apical shape of the initiating cells, but also forces the lateral membrane of the adjacent cells to bend towards the initiating cells, thereby facilitating tissue deformation. Our data thus establish a direct link between modification of epithelial polarity and initiation of epithelial folding.

The anterior and posterior dorsal transverse folds, or the dorsal folds, are epithelial folds that form on the dorsal side of the gastrulating *Drosophila* embryo at stereotypical locations coincident with the second and fifth stripes of the Runt expression (Fig. 1a–f, Supplementary Movie 1 and Supplementary Fig. 1a). Whereas the anterior fold is eventually shallow and the posterior fold deep, the initial cell shape changes are similar in both and the underlying mechanisms appear to be cell-autonomous (Supplementary Movies 2, 3 and Supplementary Fig. 1b, c).

We monitored cell shape changes using two-photon laser scanning microscopy in live embryos that express a membrane marker conjugated with the green fluorescent protein (Resille–GFP, also known as P[PTT-un1]CG8668^{117–2}). Optical sectioning of embryos at the mid-sagittal plane reveals that two stripes of dorsal cells, each three to seven cells wide, narrow their apices and shorten cell length during early gastrulation, producing two clefts on the dorsal surface that represent the first morphological signs of dorsal fold formation (Supplementary Fig. 2a and Supplementary Movie 4, see also Fig. 4b for measurements of shortening). Cells that undergo apical narrowing retain dome-like apices (Supplementary Fig. 2b), contrasting with the flattened apical surface caused by apical constriction during *Drosophila* ventral furrow formation².

We sought to identify dynamic cellular processes that precede cell shape changes. Unlike the canonical mode of epithelial folding in which spatially restricted activation of the molecular motor, myosin II (encoded by *spaghetti squash*), drives localized apical constriction to initiate tissue deformation^{1,3}, the basal levels of apical myosin remain

low and constant across the dorsal epithelium throughout the course of dorsal fold initiation with infrequent bursts of myosin activity that do not differ between the initiating and neighbouring cells (Supplementary Movie 5 and Supplementary Fig. 3a, b). These results indicate that the initiation of dorsal fold formation is not associated with differential myosin contractility.

In contrast, E-Cadherin (encoded by *shotgun*), the core component of adherens junctions, shows a cell-type-specific change in its positioning: in the initiating cells, junctions shift basally from the subapical regions where they are originally assembled, whereas in the neighbouring cells junctions maintain their original subapical positioning (Fig. 1g, Supplementary Movies 6 and 7 and Supplementary Fig. 3c). Simultaneous imaging of E-Cadherin–GFP and Resille–GFP reveals that basal shift of junctions can be observed as early as 300 s before the onset of gastrulation during the last phase of cellularization, which precedes the apical narrowing and cell shortening that occur 100–200 s after the onset of gastrulation (Fig. 1h and Supplementary Movie 8). During this seven-minute interval, junctions in the initiating cells shift approximately 10 μm basally to lie at $34 \pm 5\%$ ($n = 18$) below the apical surface, whereas junctions in the neighbouring cells show only a slight shift ($\sim 3 \mu\text{m}$) to lie at $15 \pm 4\%$ ($n = 27$) below the apical surface (Supplementary Fig. 4). The basal shift of junctions in the initiating cells increases the asymmetry in the junctional positioning on the opposite sides of the neighbouring cells that immediately flank the initiating cells. The lateral membrane of these cells becomes increasingly curved, correlating with the increased junctional asymmetry (Supplementary Fig. 5).

If the apparent basal shift of E-Cadherin positioning reflects an actual movement or remodelling of the junctions, it should be associated with an increase in the volume and surface area above the junctions. To test this hypothesis, we measured the two-dimensional parameters of area and perimeter of the apical domain in the living embryos. As the junctions shift basally in the initiating cells, both of these parameters increase, consistent with a basal movement of the junctions within the cells (Supplementary Fig. 6). We corroborated these observations by developing computer software that reconstructs and quantifies three-dimensional cell shape in fixed embryos (Fig. 1i). As the cell length increases during the last phase of cellularization, the length, volume and surface area of the apical domain in the initiating cells all increase significantly more than they do in the neighbouring cells (Fig. 1j–l), indicating that the junctional shift is accompanied by an expansion of the apical domain and that mobility of the E-Cadherin complex underlies the apparent basal shift of the junctions.

Adherens junctions are positioned to the subapical regions of the polarized epithelial cells by the concerted action of the scaffolding protein Par-3 (encoded by *bazooka* in *Drosophila*), the atypical protein kinase C (aPKC) and the MARK family kinase Par-1: apically localized aPKC and basal-laterally localized Par-1 restrict Par-3 to the subapical regions, where it directs junctional assembly^{4–10}. We found that the

¹Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA. ²Howard Hughes Medical Institute, Princeton University, Princeton, New Jersey 08544, USA. ³Department of Computer Science, Princeton University, Princeton, New Jersey 08540, USA. ⁴Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, USA. [†]Present addresses: Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA (Z.K.); Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe University, Frankfurt am Main 60483, Germany (M.K.).

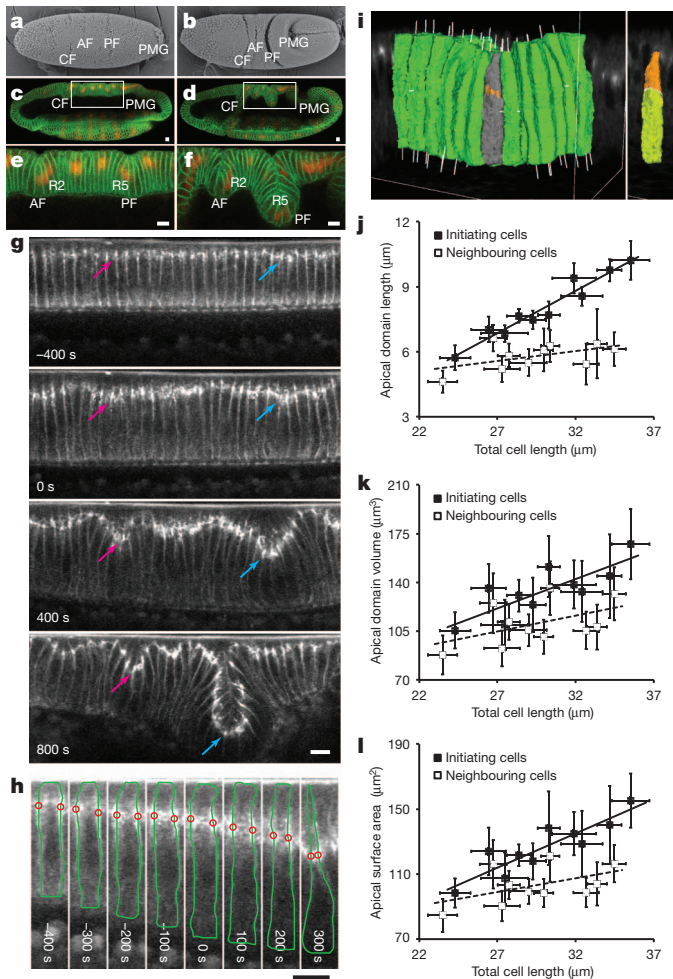


Figure 1 | Morphology and cellular dynamics during dorsal fold formation. **a, b**, Scanning electron micrographs of the dorsal surface in an early (**a**) and a late (**b**) *Drosophila* gastrula. **c–f**, Confocal mid-sagittal sections of Neurotactin (green) and Runt (red) immunofluorescence in an early (**c**) and a late (**d**) gastrula. **e, f**, A magnified view of the highlighted areas in **c** and **d**. AF, anterior fold; CF, cephalic furrow; PF, posterior fold; PMG, posterior midgut; R2, second stripe of Runt; R5, fifth stripe of Runt. **g**, Two-photon time-lapse mid-sagittal section of E-Cadherin-GFP. Arrows, junctions of initiating cells in the anterior (pink) and posterior (cyan) folds undergo basal shift. **h**, Two-photon time-lapse mid-sagittal section of E-Cadherin-GFP and Resilin-GFP in a posterior fold-initiating cell with manual traces of cell outlines (green) and junctional position (red). Scale bars, 10 μm . **i**, Three-dimensional rendering of a posterior fold based on Neurotactin immunofluorescence with an initiating cell highlighted in grey (left panel, orange patches depict thresholded Bazooka staining). The Bazooka staining is used to subdivide the cell into the apical (orange) and basal (yellow) domains (right panel). **j–l**, Scatter plots of the average length, volume and surface area of apical domain against the average total cell length in a series of fixed late cellularizing embryos with solid and dashed trend lines for the initiating and neighbouring cells. Error bars indicate s.d.

levels of Bazooka and aPKC are not differentially regulated across the dorsal epithelium and thus do not account for the observed junctional shift (Fig. 2a, Supplementary Movie 9 and Supplementary Fig. 7). In contrast, live imaging of Par-1-GFP shows that the levels of Par-1 in the presumptive initiating cells, although initially similar ($\sim 95\%$) to those in the neighbouring cells before the onset of junctional shift, reduce progressively during the last phase of cellularization to reach approximately 80% of its levels in the neighbouring cells as gastrulation commences (Fig. 2b, c and Supplementary Movie 10, $n = 7$). This differential modulation of Par-1 levels seems to require the anterior–posterior patterning system (Supplementary Fig. 8 and Supplementary Movie 11). To ask whether the reduction in Par-1 levels in the initiating

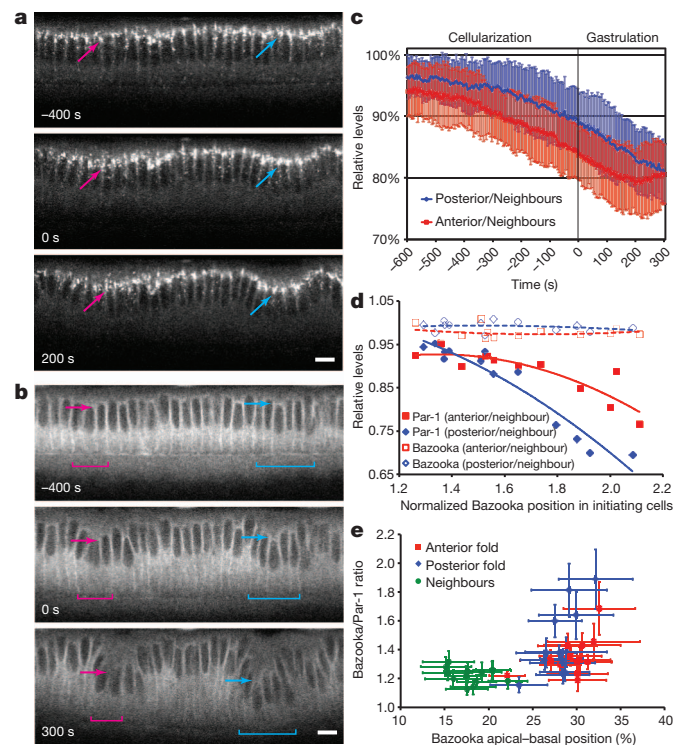


Figure 2 | The dynamics of Bazooka and Par-1 during dorsal fold initiation. **a, b**, Two-photon time-lapse mid-sagittal section of Bazooka-GFP (**a**) or Par-1-GFP (**b**). Arrows and brackets, anterior (pink) and posterior (cyan) fold-initiating cells. Scale bars, 10 μm . **c**, A time-course analysis of Par-1-GFP levels in the initiating cells relative to those in the neighbouring cells ($n = 7$). **d**, A scatter plot of the average Bazooka positioning in the initiating cells normalized by that in the neighbouring cells against the average levels of Bazooka or Par-1 relative to their respective levels in the neighbouring cells with the corresponding trend lines. **e**, A scatter plot of the average Bazooka positioning along the apical–basal axis against the average Bazooka/Par-1 ratio within individual cells. Error bars indicate s.d.

cells correlate temporally with the junctional shift, we quantified the levels of Par-1 in fixed embryos and determined the position of junctions using Bazooka staining. As Bazooka becomes more basally positioned in the initiating cells, their Par-1 levels also become lower, whereas the Bazooka levels remain constant (Fig. 2d). These analyses confirm our live imaging data and establish a correlation between the position of junctions and the ratio of Bazooka/Par-1 (Fig. 2e).

This correlation suggests that Par-1 downregulation allows Bazooka to gradually localize more basally, which in turn directs basal repositioning of junctions. To test this hypothesis, we altered the levels of Bazooka and Par-1 to investigate the function of junctional positioning during the formation of dorsal folds. Uniform reduction of Bazooka activity by RNA interference (RNAi) causes accumulation of E-Cadherin-GFP at the edges between apical and lateral surfaces, resulting in an extreme apical positioning of junctions across the epithelium (Fig. 3a and Supplementary Movie 13), similar to embryos produced by the germline clones of a strong loss-of-function allele of *bazooka* (Supplementary Fig. 9). Conversely, in *par-1* RNAi embryos, junctions are located in the lateral regions of all dorsal cells at an average position of $39 \pm 8\%$ below the surface, slightly more basal than the junctions in the initiating cells in the wild-type (Fig. 3b and Supplementary Movie 14, 30 cells from 3 embryos). Importantly, in both *bazooka* and *par-1* RNAi embryos, the junctional positioning is uniform across the entire dorsal epithelium and in each case, the initiation of dorsal folds is abolished despite the normal appearance of junction and epithelial structure (75% for *bazooka* RNAi, $n = 8$; 70% for *par-1* RNAi, $n = 10$). Thus, dorsal fold formation seems to

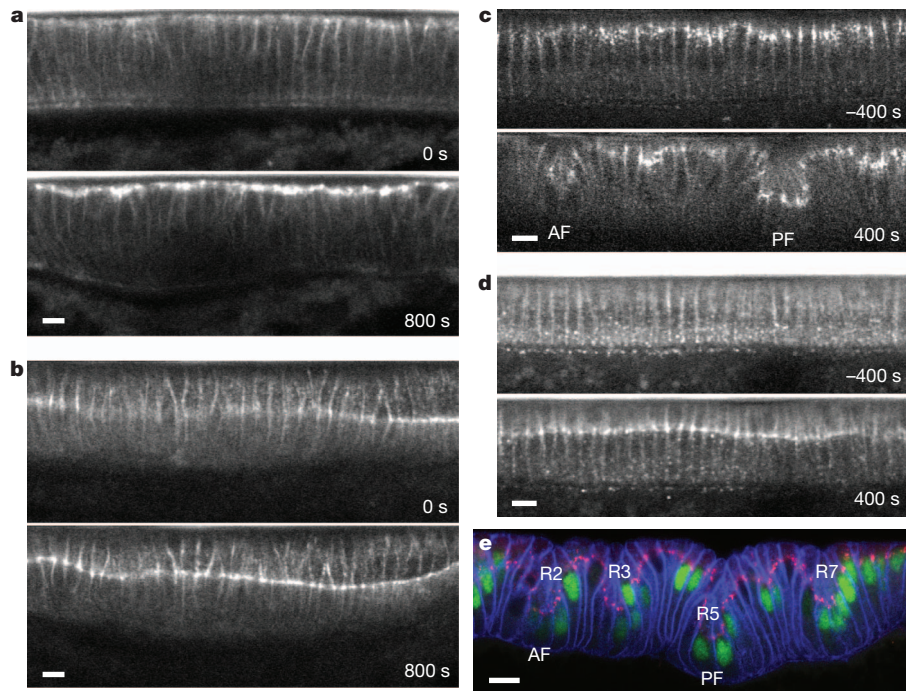


Figure 3 | Differential positioning of adherens junctions is necessary and overexpression of Bazooka can be sufficient for ectopic dorsal fold initiation. **a–d**, Two-photon time-lapse mid-sagittal section of E-Cadherin-GFP in a *bazooka* (**a**) or *par-1* (**b**) RNAi embryo or Bazooka^{S151A, S1085A}-GFP in an embryo in which the endogenous Bazooka is present (**c**) or downregulated

by *bazooka* 5'UTR RNAi (**d**). **e**, Confocal mid-sagittal section of Bazooka (red), Neurotactin (blue) and Runt (green) immunofluorescence in a Bazooka-GFP overexpression embryo. R2, R3, R5 and R7 denote the Runt stripes. AF, anterior fold; PF, posterior fold. Scale bars, 10 μ m.

require a differential positioning of junctions between the initiating cells and their neighbours.

Par-1 phosphorylates and thereby excludes Bazooka from the basal-lateral regions of a polarized epithelial cell. We examined the behaviour of Bazooka^{S151A, S1085A}, a mutant form of Bazooka that cannot be phosphorylated by the Par-1 kinase⁴. When the endogenous Bazooka is present, the GFP-tagged Bazooka^{S151A, S1085A} shows a subapical (junctional) distribution similar to the GFP-tagged wild-type form (Fig. 3c and Supplementary Movie 15). However, when we knocked down the endogenous Bazooka using RNAi, Bazooka^{S151A, S1085A} initially shows a broad distribution along the apical–basal axis and eventually coalesces in the lateral regions of all dorsal cells (Fig. 3d and Supplementary Movie 16). A similar localization was observed for wild-type Bazooka-GFP in *par-1* RNAi embryos (Supplementary Movie 17) and in both cases, dorsal fold formation is blocked. These results indicate that serine 151 and 1085 of Bazooka are two main substrates of Par-1 during dorsal fold initiation, whose differential phosphorylation determines the heterogeneous positioning of Bazooka across the dorsal epithelium and is critical for dorsal fold initiation.

When we altered the ratio of Bazooka/Par-1 by a uniform increase in Bazooka levels throughout the epithelium, we saw shifts of junctions that lead to eventual formation of epithelial folds in regions that are outside the sites of anterior and posterior folds and typically near the third and seventh stripes of Runt expression (Fig. 3e, Supplementary Movie 18 and Supplementary Fig. 10). An increase in the Bazooka levels alone in the cells that normally maintain a subapical positioning of junctions can thus be sufficient to drive junctional shift and epithelial folding, presumably by exploiting subtle local heterogeneities in Par-1 levels (Supplementary Fig. 11).

In most epithelia, aPKC phosphorylates Bazooka and becomes segregated to establish the apical domain above the junctions^{7,9,10}. We asked whether aPKC plays a role during junctional repositioning. In embryos that lack aPKC activity, the basal margin of the junctions

shows its characteristic basal shift in the initiating cells, but the apical margin unexpectedly maintains its typical subapical positioning, leading to an abnormally wide junctional domain. In contrast, the width and positioning of the junctions in the neighbouring cells appear normal (Fig. 4a and Supplementary Movie 19). These results indicate that aPKC controls the apical margin to maintain the size of the junctions, but is not required for the basal shift of junctions. These observations also decouple the junctional shift from an increase in the size of the apical membrane. The widening of junctional expanse was also observed in embryos that overexpress Bazooka^{S980A}, a mutant form of Bazooka that cannot be phosphorylated by aPKC (Supplementary Fig. 12)⁹. It seems that the segregation of aPKC from Bazooka establishes the apical domain, enabling junctional disassembly at the apical margin of the junctions.

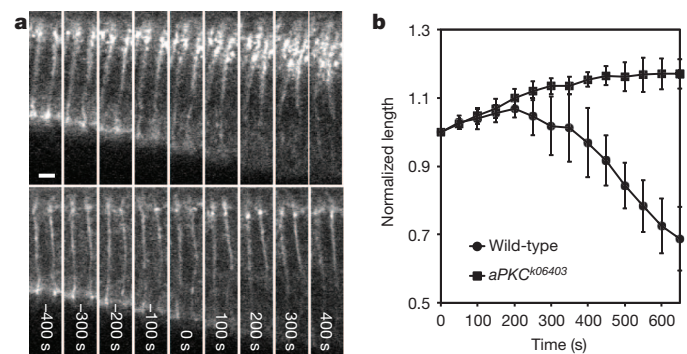


Figure 4 | Loss of aPKC results in an expansion of the junctional domain and a failure to shorten the initiating cells. **a**, Two-photon time-lapse images of E-Cadherin-GFP in an initiating (top) and a neighbouring cell (bottom) in an *aPKC* mutant embryo. Scale bar, 5 μ m. **b**, A time-course analysis of normalized cell length of the initiating cells in the wild-type ($n = 5$) and *aPKC* mutant ($n = 4$) embryos. Error bars indicate s.d.

Although the basal shift of junctions occurs in the initiating cells in the *aPKC* mutant embryo, these cells fail to shorten and the dorsal folds do not form properly (Fig. 4b and Supplementary Movie 19). It seems that as the basal margin of the junctions shifts basally in the initiating cells in response to a decrease in Par-1 levels, their apical margin needs to become disassembled in an *aPKC*-dependent manner so that the subsequent apical cell shape changes could occur.

In this report, we present evidence that dorsal fold initiation requires the establishment of distinct ratios of Bazooka/Par-1 that impose different positions for the adherens junctions in the initiating cells and their neighbours. We propose that the differential positioning of junctions facilitates epithelial folding through two cellular processes (Supplementary Fig. 13). Within the initiating cells, the resultant increase in the non-adherent apical surface after junctional shift may be unstable such that a shrinkage of the apical domain is triggered to restore the balance between cell surface tension and local adhesive forces¹¹. The shortened cells thus produced would then create a localized structural inhomogeneity in the epithelium where buckling would preferentially occur. Second, in the immediate flanking cells, a junctional asymmetry is produced because the basal positioning in the initiating cells on one side and the subapical positioning in the neighbouring cells on the other must be accommodated. Because all junctions in an epithelium are mechanically coupled¹², the asymmetry may cause the lateral surfaces to curve and cells to bend towards the shortened initiating cells. This bending would drive and deepen any buckles or folds initiated in the epithelial sheet.

Directional movement of the cadherin complex along the apical-basal axis has been observed previously in cultured cells *in vitro*¹³ but, to our knowledge, *Drosophila* gastrulation provides the first case where such movement has been described in an intact developing organism. When the shifts occur in stripes as they do on the dorsal side of the *Drosophila* embryo, they seem to initiate infolding of the epithelium. In tissues in which the levels of cortical myosin are low and constant, junctional repositioning regulated by Par-1/Bazooka interactions may play a more prominent role in epithelial folding than does differential activation of cortical contractility. Junctional repositioning may also represent an important mechanism in folding events that do not lead to internalization or delamination, or where the integrity of junctions within the epithelia must be maintained. How junctions are repositioned while maintaining junctional integrity is unclear, but in principle the process could involve remodelling via local endocytic trafficking, or lateral movement of the intact junctions in the membrane¹⁴. Regardless of the mechanism, dorsal fold formation represents an emergent model in which the insights into this alternative mode of epithelial folding could be further analysed.

METHODS SUMMARY

Detailed information about reagents and methods used in this paper, including the *Drosophila* stocks, RNAi, live imaging, immunofluorescence, scanning electron microscopy, image processing, three-dimensional cell boundary reconstruction and image quantification is described in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 May 2011; accepted 9 February 2012.

Published online 28 March 2012.

1. Sawyer, J. M. *et al.* Apical constriction: a cell shape change that can drive morphogenesis. *Dev. Biol.* **341**, 5–19 (2010).
2. Sweeton, D., Parks, S., Costa, M. & Wieschaus, E. Gastrulation in *Drosophila*: the formation of the ventral furrow and posterior midgut invaginations. *Development* **112**, 775–789 (1991).
3. Martin, A. C., Kaschube, M. & Wieschaus, E. F. Pulsed contractions of an actin-myosin network drive apical constriction. *Nature* **457**, 495–499 (2009).
4. Benton, R. & St Johnston, D. *Drosophila* PAR-1 and 14-3-3 inhibit Bazooka/PAR-3 to establish complementary cortical domains in polarized cells. *Cell* **115**, 691–704 (2003).
5. Harris, T. J. & Peifer, M. Adherens junction-dependent and -independent steps in the establishment of epithelial cell polarity in *Drosophila*. *J. Cell Biol.* **167**, 135–147 (2004).
6. Harris, T. J. & Peifer, M. The positioning and segregation of apical cues during epithelial polarity establishment in *Drosophila*. *J. Cell Biol.* **170**, 813–823 (2005).
7. Krahn, M. P., Buckers, J., Kastrop, L. & Wodarz, A. Formation of a Bazooka–Stardust complex is essential for plasma membrane polarity in epithelia. *J. Cell Biol.* **190**, 751–760 (2010).
8. McGill, M. A., McKinley, R. F. & Harris, T. J. Independent cadherin–catenin and Bazooka clusters interact to assemble adherens junctions. *J. Cell Biol.* **185**, 787–796 (2009).
9. Morais-de-Sá, E., Mirouse, V. & St Johnston, D. *aPKC* phosphorylation of Bazooka defines the apical/lateral border in *Drosophila* epithelial cells. *Cell* **141**, 509–523 (2010).
10. Walther, R. F. & Pichaud, F. Crumbs/DaPKC-dependent apical exclusion of Bazooka promotes photoreceptor polarity remodeling. *Curr. Biol.* **20**, 1065–1074 (2010).
11. Lecuit, T. & Lenne, P. F. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature Rev. Mol. Cell Biol.* **8**, 633–644 (2007).
12. Martin, A. C., Gelbart, M., Fernandez-Gonzalez, R., Kaschube, M. & Wieschaus, E. F. Integration of contractile forces during tissue invagination. *J. Cell Biol.* **188**, 735–749 (2010).
13. Kametani, Y. & Takeichi, M. Basal-to-apical cadherin flow at cell junctions. *Nature Cell Biol.* **9**, 92–98 (2007).
14. Cavey, M. & Lecuit, T. Molecular bases of cell-cell junctions stability and dynamics. *Cold Spring Harb. Perspect. Biol.* **1**, a002998 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Martin, J. McDonald, D. St Johnston and J. Zallen for providing flies and antibodies; J. Goodhouse and S. Thiberge for assistance in microscopy; G. Deshpande, members of the Wieschaus and Schüpbach laboratories for helpful comments on the manuscript and discussion. This work is supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation to Y.-C.W., a National Institutes of Health/National Institute of General Medical Sciences P50 grant (GM071508) to Z.K. and M.K., and a National Institute of Child Health and Human Development grant (5R37HD15587) to E.F.W. E.F.W. is an investigator of the Howard Hughes Medical Institute.

Author Contributions Y.-C.W. conceived the project, performed all experiments and analysed the data, except those experiments that involve scanning electron microscopy, which were performed by E.F.W.; Z.K. and M.K. developed the software for three-dimensional reconstruction. Y.-C.W. and E.F.W. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.F.W. (efw@princeton.edu).

METHODS

Drosophila genetics. *Drosophila* stocks used for live imaging were: *Resille-GFP*¹⁵, *myosin-GFP* (*Spaghetti squash-GFP*)¹⁶, *membrane-mCherry*¹² (also known as *P[sqh-mCherry.membrane]*), *E-Cadherin-GFP*¹⁷, *Par-1 protein trap*¹⁸ (also known as *P[PTT-GC]par-1*^{CCO1981}), *mat-tub-GFP-Par-1* (ref. 19). *UASp-Bazooka-GFP*²⁰ was driven maternally by one copy (67C) of the *matzTub-Gal4VP16* driver in live imaging experiments and one or two copies (67C; 15) in overexpression experiments. *UASp-Bazooka*^{S151A, S1085A-GFP} and *UASp-Bazooka*^{S980A-GFP} were driven by *matzTub-Gal4VP16* (67C; 15). Mutant stocks used were: *run1*^{LB5}, *torso*^{XR/torso}^{PM51}, *torso-like*^{1/torso-like}⁴, *bicoid*^{E1}, *nanos*^{BN}, *torso-like*³. Germline clones of *bazooka*^{GD21} and *aPKC*^{K06402} were generated using the FLP-recombinase/dominant female sterile system with the *ovo*^{D2} *FRT*^{19A} or *FRT*^{G13} *ovo*^{D1} chromosome (*FRT*^{19A} and *FRT*^{G13} are also known as *P[neoFRT]19A* and *P[FRT(w^h)]G13*).

RNAi. Double-stranded RNAs were synthesized using a MEGAscript T7 kit (Ambion) from PCR products that contain the T7 promoter sequence (5'-TAATACGACTCACTATAGGTTACT-3') at each end. The PCR products used in *in vitro* transcription reactions were amplified from 0–4 h embryonic cDNA using the following primer pairs: *bazooka*, 5'-GACGTTTCTTGCTAAGCGG-3', 5'-TTTCGACAGTGTAGGTCCAAA-3'; *bazooka* 5'UTR (for knockdown of endogenous but not transgenic *bazooka*), 5'-AATGCGCGCGTGTGAATCA CAC-3', 5'-ACGACCGCATCATCATCGTCG-3'; *par-1*, 5'-CACGTTCTG CCGTAGCC-3', 5'-GCTTGGGATCGGCTAAATC-3. Double-stranded RNAs were injected into the embryos during the syncytial blastoderm stage, typically 3–4 h before imaging.

Live imaging, immunofluorescence and scanning electron microscopy. Two-photon live embryo imaging was performed on a custom-made system built on an upright Olympus BX51 microscope that is equipped with a Ti:sapphire tunable laser ranged from 720 to 960 nm (Coherent). Single-photon confocal imaging was performed on a Leica SP5 system. Immunofluorescence was performed on heat-methanol fixed embryos²¹. Antibodies used were mouse monoclonal anti-Neurotactin (BP106, Developmental Studies Hybridoma Bank, 1:20), rabbit anti-Runt (1:1,000), rabbit anti-Armadillo (1:200), rabbit anti-Par-1 (ref. 22, 1:500), rabbit anti-PKC ζ C20 (1:1,000, Santa Cruz Biotechnology), and guinea pig anti-Bazooka²³ (1:500) and were visualized by Alexa 488-, 568- and 647-conjugated secondary antibodies (Molecular Probes). Scanning electron microscopy was performed on a Hitachi TM-1000 system as previously described¹². Images were processed, assembled into figures and converted into movies using ImageJ, Adobe Photoshop and Adobe Illustrator.

Three-dimensional image processing and cell boundary reconstruction. The algorithm for three-dimensional reconstruction and analysis was implemented in C++ using the Qt library and OpenGL for the graphical user interface. Libtiff was used for loading image stacks. Image stacks of three-channel immunofluorescence were used. All channels of the 8-bit image volume were initially scaled to a 1:1:1 aspect ratio (voxel size, 0.16 $\mu\text{m} \times 0.16 \mu\text{m} \times 0.16 \mu\text{m}$) and down-sampled by 80% (voxel size, 0.20 $\mu\text{m} \times 0.20 \mu\text{m} \times 0.20 \mu\text{m}$) to reduce image noise. The contrast of Neurotactin immunofluorescence was enhanced by adaptive histogram adjustment using eight equally spaced histograms along the z-dimension and the 10th percentile as the minimum intensity value for intensity adjustment. Two passes of rank filtering were used to fill weak regions of Neurotactin staining, closing holes between cells: a rank filter using the 95th percentile intensity value in a sphere with radius 0.6 μm , followed by a rank filter using the 10th percentile intensity value in a sphere with radius 0.8 μm . Edges were found using a difference of Gaussians (DOG) approximation to a three-dimensional Marr–Hildreth edge detector where the zero-crossing was positioned at low threshold of 4 and high threshold of 30 to generate two binary image volumes²⁴. A ‘rolling ball’ algorithm applied to the high-threshold volume was used to repair holes in the epithelium due to ongoing cellularization. The algorithm was computed efficiently using boundaries in a Euclidian distance transform (EDT)²⁵. Briefly, a boundary at distance of 3 μm was defined using the EDT. A second boundary, also at 3 μm from the first boundary, was used to approximate the result of rolling a sphere on the high-threshold binary image generated from the Marr–Hildreth operator. The repaired boundary was applied to the low-threshold binary image to obtain a binary image where the outer boundary of the epithelium was repaired. This binary image was then thinned by three-dimensional surface thinning²⁶. Connected components in the surface-thinned binary image were found by depth first search. Components of fewer than 100 voxels were removed as noise. Cells were found in this image by hierarchical application of a seeded Watershed algorithm^{27,28}. Seed regions were defined hierarchically by gradually applying an increasing threshold to an EDT of the thinned binary image until a Watershed-segmented region reached a volume of less than 640 μm^3 . Regions smaller than 40 μm^3 were removed as noise. The segmented cell regions were then converted into three-dimensional triangle meshes by the Marching Cubes algorithm²⁹. Lastly, the resulting meshes were

adapted to the intensity of the image by a finite difference approximation to an Active Surface³⁰.

Image quantification. The frequency of myosin bursts (Supplementary Fig. 3b) was measured in confocal time-lapse images of myosin-GFP and membrane-mCherry. These data sets have a z resolution of 1 μm and cover a 10 μm distance from the apical cortex with a temporal resolution ranging between 15 and 22 s per frame. The intense myosin structures were visually identified from each z slice of the image stacks. The total number of myosin bursts was the sum of myosin bursts from all z slices throughout the duration of imaging. The frequency was then calculated by dividing the total number of bursts by the imaging duration and the number of cells in which the bursts were counted. The imaging duration ranges between 436 and 689 s. The numbers of initiating and neighbouring cells that were counted range between 30 to 41 and 51 to 86, respectively.

Time-course analyses of junctional positioning and apical domain size (Supplementary Fig. 4, 6) were performed using two-photon time-lapse images of E-Cadherin-GFP and Resille-GFP. The onset of gastrulation was defined by the onset of anterior cell movement driven by the posterior midgut invagination. For the analysis of junctional positioning, the central initiating cell of the anterior and posterior folds and a third cell that resides in the region between the anterior and posterior fold that shows minimal junctional movement were chosen to represent the initiating and neighbouring cells. The vertical distance between the visually defined centre of the junctional complex and the apex of the cell was measured in Image J to represent the positioning of the junctions. For the analysis of apical domain size, the central initiating cell of the posterior fold and a representative neighbouring cell in the region between the anterior and posterior folds were chosen. The apical domain above the junctions was manually outlined based on the membrane fluorescence of Resille-GFP and measured for its area and perimeter in ImageJ.

Correlation analysis between differential junctional displacement and lateral membrane curvature (Supplementary Fig. 5) was performed using two-photon time-lapse images of E-Cadherin-GFP. Cells that are in the immediate flanking regions of the initiating cells and show a marked asymmetry of junctional positioning on the opposite sides were chosen for these measurements. The differential junctional displacement, which defines the extent of junctional asymmetry, was calculated by subtracting the length of the apical domain on the distal side (y) from that on the proximal side (x). The lateral membrane curvature was defined as the ratio between the height (h) and the chord (C) of the arc of lateral membrane on the distal side of the cell.

Three-dimensional cell shape measurements (Fig. 1j–l) were made in image stacks of late cellularization embryos that have been stained for Bazooka and Neurotactin and processed for three-dimensional reconstruction as described above. The position of Bazooka was defined by the ‘Bazooka junctional triangles’. Briefly, an average intensity of Bazooka was first assigned for the voxels that intersect with a three-dimensional triangle mesh in the reconstructed cell boundary. The Bazooka junctional triangles were then selected based on an intensity threshold of the 99th percentile of the Bazooka intensity histogram. For each of the centroids of the Bazooka junctional triangles, a three-dimensional principal component analysis (PCA) was performed to determine the Bazooka mean position (a point on a plane) and the eigenvector corresponding to smallest eigenvalue (plane normal). These were then used to define the Bazooka junctional plane that subdivides the cell into the apical and basal domains. The geometric measurements were performed as follows: three-dimensional PCA was applied to all of the vertices of the triangle mesh and the long direction of the cell was defined using the eigenvector corresponding to the largest eigenvalue. The apical domain length was measured by first creating vectors between the cell centroid and each mesh vertex on the apical side of the Bazooka junctional plane. These vectors were then projected onto the long direction vector of the cell. The length of the longest projected vector was used as the apical domain length. The basal domain length was measured similarly, using triangle mesh vertices on the basal domain of the cell. The total cell length was computed by a sum of the apical and basal length. The apical volume was computed by voxelizing the three-dimensional triangle meshes, and summing the volumes of voxels apical to the Bazooka junctional plane. Similarly, the apical surface area was computed by summing the areas of the mesh triangles for which the triangle centroid falls on the apical side of the Bazooka junctional plane. The initiating cells were selected on the basis of their location and junctional positioning. Approximately 500 dorsal cells in the region between the first and the seventh stripe of Runt were ranked by the apical domain length and the top 150 cells were selected for further analysis. A second selection was performed to isolate those that are in close proximity to the second and fifth stripes of Runt. Of these cells, those whose apical domain length was above the average were used for analysis. For the early stage embryos that showed no junctional shift, only the location-based selection was made. Cells that reside in the region between the anterior and posterior folds with junctional positioning that was below average were used as the neighbouring cells.

Time-course analysis of Par-1 dynamics (Fig. 2c) was performed using two-photon time-lapse images of Par-1-GFP. The average fluorescent intensity of Par-1-GFP in manually selected areas consisting of two anterior or four posterior initiating cells was measured and normalized by that in areas consisting of four neighbouring cells that reside in the regions between the anterior and posterior folds. The onset of gastrulation was defined by the onset of anterior cell movement driven by posterior midgut invagination.

Bazooka and Par-1 immunofluorescence (Fig. 2d, e and Supplementary Fig. 11) was quantified in image stacks of fixed embryos that were triply labelled for Bazooka, Par-1 and Neurotactin and processed for three-dimensional reconstruction as described above to define the Bazooka junctional triangles. Junctional intensity of Bazooka within a cell was measured within and normalized by the Bazooka junctional volume that was defined by the voxelization of Bazooka junctional triangles. The basal-lateral intensity of Par-1 in each cell was measured within and normalized by the volume within a two-voxel distance from the cell boundary basal to the Bazooka junctional plane. For Fig. 2d, e, the anterior and posterior fold-initiating cells were selected on the basis of location and above-average Bazooka positioning, whereas the neighbouring cells were selected from the cells that reside in the region between the anterior and posterior folds with below-average Bazooka positioning. For Supplementary Fig. 11, the wild-type and Bazooka overexpression embryos were fixed, stained and imaged in parallel under identical conditions. Cell selection was performed as in Fig. 2d, e.

Time-course of cell shortening in the initiating cells (Fig. 4b) was analysed by measuring the total cell length of five and four initiating cells each from wild-type and *aPKC* mutant embryos in ImageJ. The measurements were normalized by cell length at time zero, which was defined by the onset of anterior cell movement driven by the posterior midgut invagination.

15. Morin, X., Daneman, R., Zavortink, M. & Chia, W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 15050–15055 (2001).
16. Royou, A., Sullivan, W. & Karess, R. Cortical recruitment of nonmuscle myosin II in early syncytial *Drosophila* embryos: its role in nuclear axial expansion and its regulation by Cdc2 activity. *J. Cell Biol.* **158**, 127–137 (2002).
17. Oda, H. & Tsukita, S. Real-time imaging of cell-cell adherens junctions reveals that *Drosophila* mesoderm invagination begins with two phases of apical constriction of cells. *J. Cell Sci.* **114**, 493–501 (2001).
18. Lighthouse, D. V., Buszczak, M. & Spradling, A. C. New components of the *Drosophila* fusome suggest it plays novel roles in signaling and transport. *Dev. Biol.* **317**, 59–71 (2008).
19. Shulman, J. M., Benton, R. & St Johnston, D. The *Drosophila* homolog of *C. elegans* PAR-1 organizes the oocyte cytoskeleton and directs *oskar* mRNA localization to the posterior pole. *Cell* **101**, 377–388 (2000).
20. Benton, R. & St Johnston, D. A conserved oligomerization domain in *Drosophila* Bazooka/Par-3 is important for apical localization and epithelial polarity. *Curr. Biol.* **13**, 1330–1334 (2003).
21. Müller, H. A. & Wieschaus, E. *armadillo*, *bazooka*, and *stardust* are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. *J. Cell Biol.* **134**, 149–163 (1996).
22. McDonald, J. A., Khodyakova, A., Aranjuez, G., Dudley, C. & Montell, D. J. PAR-1 kinase regulates epithelial detachment and directional protrusion of migrating border cells. *Curr. Biol.* **18**, 1659–1667 (2008).
23. Simões, M. *et al.* Rho-kinase directs Bazooka/Par-3 planar polarity during *Drosophila* axis elongation. *Dev. Cell* **19**, 377–388 (2010).
24. Marr, D. & Hildreth, E. Theory of edge detection. *Proc. R. Soc. Lond. B Biol. Sci.* **207**, 187–217 (1980).
25. Maurer, C. R. Jr, Qi, R. & Raghavan, V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 265–270 (2003).
26. Palágyi, K. *et al.* A sequential 3D thinning algorithm and its medical applications information processing in medical imaging. *Lect. Notes Comput. Sci.* **2082**, 409–415 (2001).
27. Vincent, L. & Soille, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 583–598 (1991).
28. Beucher, S. & Meyer, F. In *Mathematical Morphology in Image Processing* Ch. 12 (ed. Dougherty, E. R.) 433–481 (Marcel Dekker, 1993).
29. Lorensen, W. E. & Cline, H. E. Marching cubes: A high resolution 3D surface construction algorithm. *SIGGRAPH '87: Proceedings of the 14th annual conference on computer graphics and interactive techniques* **21**, 163–169 (1987).
30. Cohen, L. D. & Cohen, I. Finite element methods for active contour models and balloons for 2D and 3D images. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 1131–1147 (1991).

Local generation of glia is a major astrocyte source in postnatal cortex

Woo-Ping Ge¹, Atsushi Miyawaki², Fred H. Gage³, Yuh Nung Jan¹ & Lily Yeh Jan¹

Glial cells constitute nearly 50% of the cells in the human brain¹. Astrocytes, which make up the largest glial population, are crucial to the regulation of synaptic connectivity during postnatal development². Because defects in astrocyte generation are associated with severe neurological disorders such as brain tumours³, it is important to understand how astrocytes are produced. Astrocytes reportedly arise from two sources^{4–6}: radial glia in the ventricular zone and progenitors in the subventricular zone, with the contribution from each region shifting with time. During the first three weeks of postnatal development, the glial cell population, which contains predominantly astrocytes, expands 6–8-fold in the rodent brain⁷. Little is known about the mechanisms underlying this expansion. Here we show that a major source of glia in the postnatal cortex in mice is the local proliferation of differentiated astrocytes. Unlike glial progenitors in the subventricular zone, differentiated astrocytes undergo symmetric division, and their progeny integrate functionally into the existing glial network as mature astrocytes that form endfeet with blood vessels, couple electrically to neighbouring astrocytes, and take up glutamate after neuronal activity.

Most radial glia have finished producing their share of astrocytes and have begun to disappear shortly after birth^{4–6}; astrocytes are therefore thought to derive mainly from progenitors in the subventricular zone (SVZ) at later stages⁸. The massive expansion of glia within the first three postnatal weeks presents a daunting task for their presumed SVZ progenitors. This task is rendered even more challenging by the thickening of the cortex compounded by the disappearance of radial glia, which provides the migratory tracks for newly formed astrocytes⁹. We used electroporation to transfect green fluorescent protein (GFP) plasmids into SVZ/radial glial cells of mice at postnatal days (P)0–2 to label them with GFP *in vivo* and to trace their progeny at P16–20 (Fig. 1a and Supplementary Fig. 1). Only a very small percentage (about 3%) of the astrocytes derived postnatally from SVZ/radial glial cells reached cortical layers I–IV; most were left behind in SVZ/white matter (75%) and layers V–VI (22%) (Fig. 1b, c). It therefore seems that huge numbers of cortical astrocytes generated postnatally might arise from a more efficient process, such as local cell proliferation (Ki67⁺; Fig. 1d), rather than from SVZ progenitors. Whereas glial cell division within the cortex was reported half a century ago, on the basis of studies involving [³H]thymidine incorporation into DNA¹⁰, the extent of the contribution of local glial division to postnatal astrocyte production remained unknown, owing to the difficulty in distinguishing glia generated locally from glia derived from other sources. In this study, we have obtained evidence to support the hypothesis that the local generation of astrocytes within the postnatal cortex is a major source of glia.

To label locally generated glia, we used a replication-defective murine leukaemia retrovirus (MLV) to express GFP in infected dividing cells and their progeny in postnatal cortex *in vivo*. This retrovirus specifically infects proliferating cells and has been used for cell-fate tracing in SVZ and the hippocampal subgranular zone (SGZ) *in vivo*^{8,11}. We injected

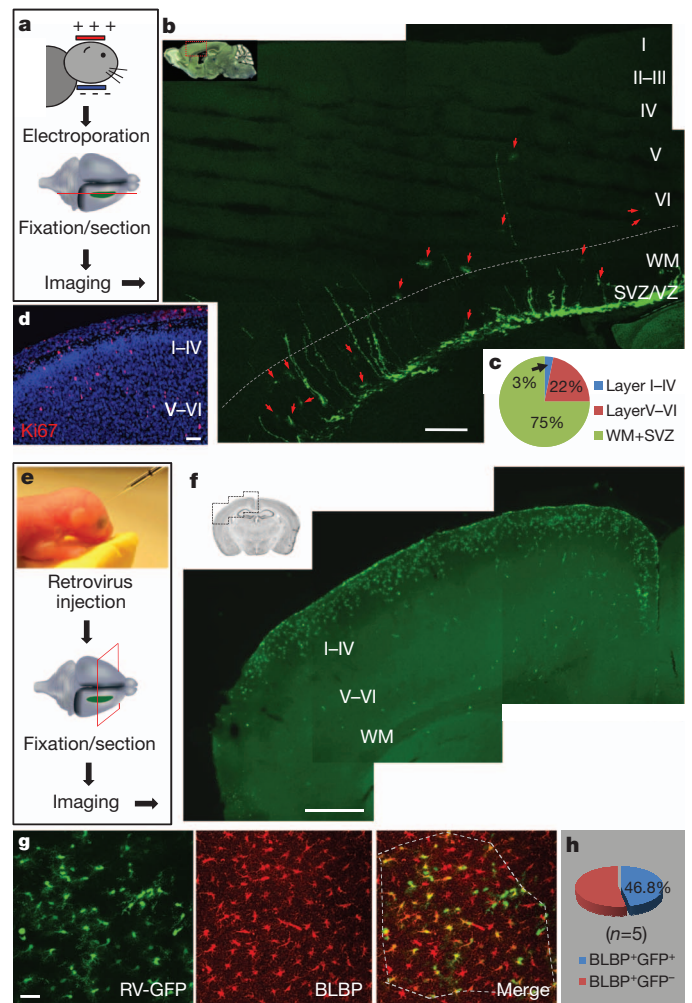


Figure 1 | Locally generated glia as a major source of astrocytes.

a, Procedure to label SVZ/radial glia-derived astrocytes by electroporation. **b**, The distribution of astrocytes (arrows) 2 weeks after electroporation. VZ, ventricular zone. **c**, Percentages of astrocytes at different locations. WM, white matter. **d**, Proliferating cells (Ki67⁺, red) in a cortical section of P3 mouse. Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI, blue). **e**, Procedure to label locally proliferating cells by retrovirus. **f**, Cells labelled by retrovirus (green). **g**, Image of infected astrocytes. Astrocytes (BLBP⁺, red) with GFP (GFP⁺BLBP⁺) or without GFP (GFP⁺BLBP⁻) in the outlined region (dashed line) were included for analysis in **h**. RV, retrovirus. **h**, Percentages of astrocytes labelled by retrovirus injected locally, calculated as 100 × (BLBP⁺GFP⁺ cells/BLBP⁺ cells). Scale bars, 200 μm (**b**), 50 μm (**d**), 500 μm (**f**) and 40 μm (**g**).

¹Howard Hughes Medical Institute, Departments of Physiology, Biochemistry and Biophysics, University of California at San Francisco, 1550 4th Street, San Francisco, California 94158, USA. ²Life Function and Dynamics, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, and Brain Science Institute, RIKEN, Wako-city, Saitama, 351-0198, Japan. ³Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, California 92037, USA.

viruses locally into layers I–IV of the barrel or motor cortex of wild-type mice at P0–6 (Fig. 1e, f) and examined GFP expression 1 week later in samples that were also stained with antibodies against brain lipid-binding protein (BLBP) (Fig. 1g), which labels radial glia during embryonic development and astrocytes in the postnatal brain¹². Whereas about 30% of infected cells were NG2 glia (27.6%, $n = 662$ infected cells, Supplementary Fig. 2), 55–70% of infected cells were astrocytes (BLBP⁺, 56.9%, $n = 369$ GFP⁺ cells, Fig. 1g; GFAP⁺, 68.6%, $n = 662$ GFP⁺ cells, Supplementary Fig. 2), indicating that these astrocytes originated locally *in vivo*.

To determine whether a major astrocyte source was derived from the local generation of glia, we injected retroviruses with higher titre ($1 \mu\text{l}$, $(1-3) \times 10^7$) into the cortex of P0–2 mice and compared the number of GFP-expressing astrocytes (BLBP⁺GFP⁺) with the total number of astrocytes (BLBP⁺) within an infected region 7–10 days

after injection. We found that nearly half of the astrocytes were doubly labelled (GFP⁺BLBP⁺, $46.8 \pm 3.8\%$, $n = 5$ mice; Fig. 1f–h). Because the half-life of infectivity of MLV retrovirus is 5–8 h at 37 °C, the doubly labelled astrocytes probably correspond to astrocytes undergoing division in the time window of 5–8 h plus the progeny they generate over the course of 7–10 days. Control studies revealed no differences in the morphology of astrocytes (Fig. 1g), the density of dividing cells (Supplementary Fig. 3) or the percentage of GFAP-occupied area (Supplementary Fig. 4) in brain regions with or without retroviral infection. Our observations therefore suggest that local proliferation is a major source of astrocytes in the postnatal cortex.

To test the possibility that multiple dividing cell types infected by retroviruses, including astrocytes, NG2 glia and perhaps some unknown progenitors in the cortex, gave rise to these GFP-expressing astrocytes, we labelled acute brain slices with the nuclear marker

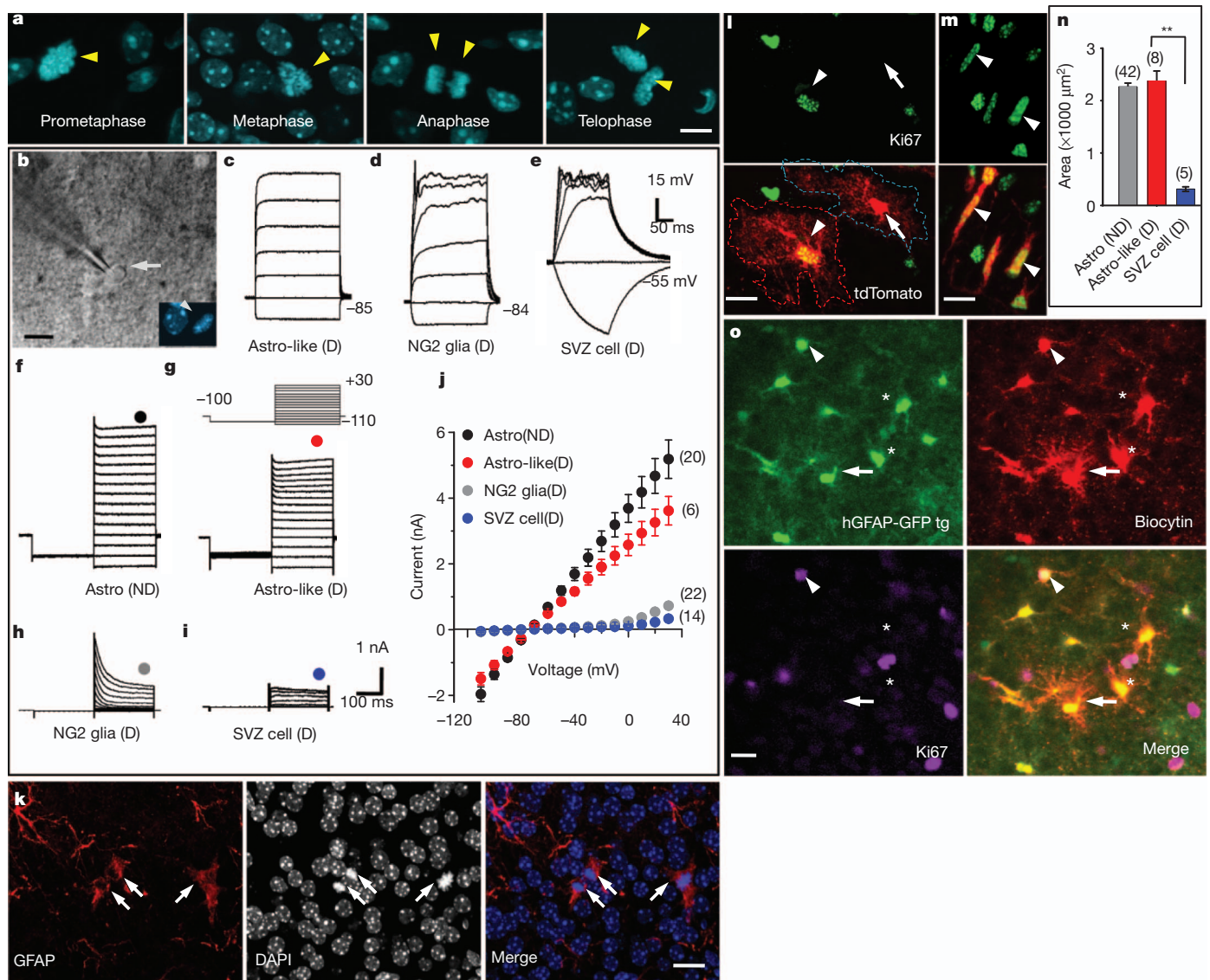


Figure 2 | Properties of dividing cells within the cortex. **a**, Nuclei (arrowheads, Hoechst 33342) of dividing cells at different mitotic stages in acute brain slices. **b**, A dividing cell (differential interference contrast, arrow) at prometaphase (Hoechst 33342, arrowhead). **c–e**, Voltage responses from an Astro-like-D cell (**c**), a dividing NG2 glia (**d**) and a dividing SVZ cell (**e**). **f–i**, Current responses from a non-dividing (ND) astrocyte (**f**), an Astro-like-D cell (**g**), a dividing NG2 glia (**h**) and a dividing SVZ cell (**i**) evoked by step voltages (inset, **g**). **j**, Current–voltage curves in **f–i** (the circles indicate the positions of measurements). **k**, Astro-like-D cells (arrows; two at telophase, one at metaphase) stained with anti-GFAP (red). **l**, **m**, Morphology of a non-dividing astrocyte (arrow) and an Astro-like-D cell (arrowhead, **l**) in the cortex, and dividing cells (arrowheads) in the SVZ (**m**) of a P8 hGFAP-CreER;Ai14 transgenic mouse. **n**, Summary of the area covered by the processes of non-dividing astrocytes (grey), Astro-like-D cells (red) and SVZ dividing cells (blue) ($10\text{-}\mu\text{m}$ z-projection with soma included, one-way analysis of variance followed by a Bonferroni post-hoc test; two asterisks, $P < 0.01$). **o**, A non-dividing astrocyte (arrow) was loaded with biocytin. A dividing astrocyte is labelled by biocytin (arrowhead). tg, transgenic. Scale bars, $5 \mu\text{m}$ (**a**) and $10 \mu\text{m}$ (**b**, **k–m**, **o**). Error bars indicate s.e.m.

Hoechst 33342, a dye that can permeate live cell membranes, to distinguish dividing cells from non-dividing cells (Fig. 2a). In some experiments we also used slices from CAG-Fucci-Green transgenic mice to identify dividing cells in the SVZ and cortex (Supplementary Fig. 5). In this line, the green fluorescent protein mAG accumulates specifically during the S (synthesis) to M (mitosis) stages of the cell cycle¹³, thus facilitating the identification of dividing cells for whole-cell patch-clamp recordings (Fig. 2b). Excluding cells of the vascular system, 94.6% (87 of 92 cells, P6–13) of the dividing cells in the cortex fell into two groups: dividing NG2 glia¹⁴, with characteristic small sodium currents and rectifying current–voltage (*I*–*V*) curve (Fig. 2d, h, j) and astrocyte-like dividing cells (Astro-like-D; Fig. 2c, g, j), so named for their similarity to differentiated astrocytes (Fig. 2f, j). Astrocytes characteristically displayed large, delayed rectifier potassium currents (K_{dr}) and large, inwardly rectifying potassium currents (K_{ir}) but no sodium currents, and they had a linear *I*–*V* curve¹⁵ (Fig. 2f, j). In contrast, dividing cells recorded in the SVZ (Supplementary Fig. 5) had no K_{ir} current and a very small K_{dr} (Fig. 2e, i, j), typical of immature progenitors¹⁶. To further characterize the Astro-like-D cells, our immunostaining revealed that they were GFAP⁺ but Nestin[−] (Fig. 2k and Supplementary Fig. 6). We then compared the morphology of

Astro-like-D cells with that of mature astrocytes or SVZ dividing progenitors in hGFAP-CreER;Ai14 transgenic mice. Crossing Ai14 transgenic mice¹⁷ with hGFAP-CreER transgenic mice¹⁸ allowed robust expression of the red fluorescent protein tdTomato after inducible astrocyte-specific Cre-mediated recombination. We administered tamoxifen to hGFAP-CreER;Ai14 transgenic mice at P0–2 and assessed the cellular morphology 1 week later (Supplementary Fig. 7). Astro-like-D cells (Ki67⁺tdTomato⁺) had a complex morphology comparable to that of neighbouring mature astrocytes (Ki67[−]tdTomato⁺; Fig. 2l, n and Supplementary Fig. 8). In contrast, SVZ dividing progenitors had a bipolar/unipolar morphology (Ki67⁺tdTomato⁺; Fig. 2m, n and Supplementary Fig. 8). Because coupling by means of gap junctions is a hallmark feature of astrocytes, we injected biocytin into individual non-dividing astrocytes in brain slices of hGFAP-GFP transgenic mice, which express GFP under the control of human astrocyte-specific GFAP promoter¹⁹, and found that Astro-like-D cells (Ki67⁺GFP⁺) and mature astrocytes (Ki67[−]GFP⁺; Fig. 2o) were coupled by means of gap junctions. Thus, unlike SVZ dividing progenitors (Fig. 2e, i, j) and glioblasts²⁰, Astro-like-D cells in the cortex are differentiated astrocytes. Taken together with a previous report²¹ and our tracing results from NG2-CreBac/ERTM;Ai14

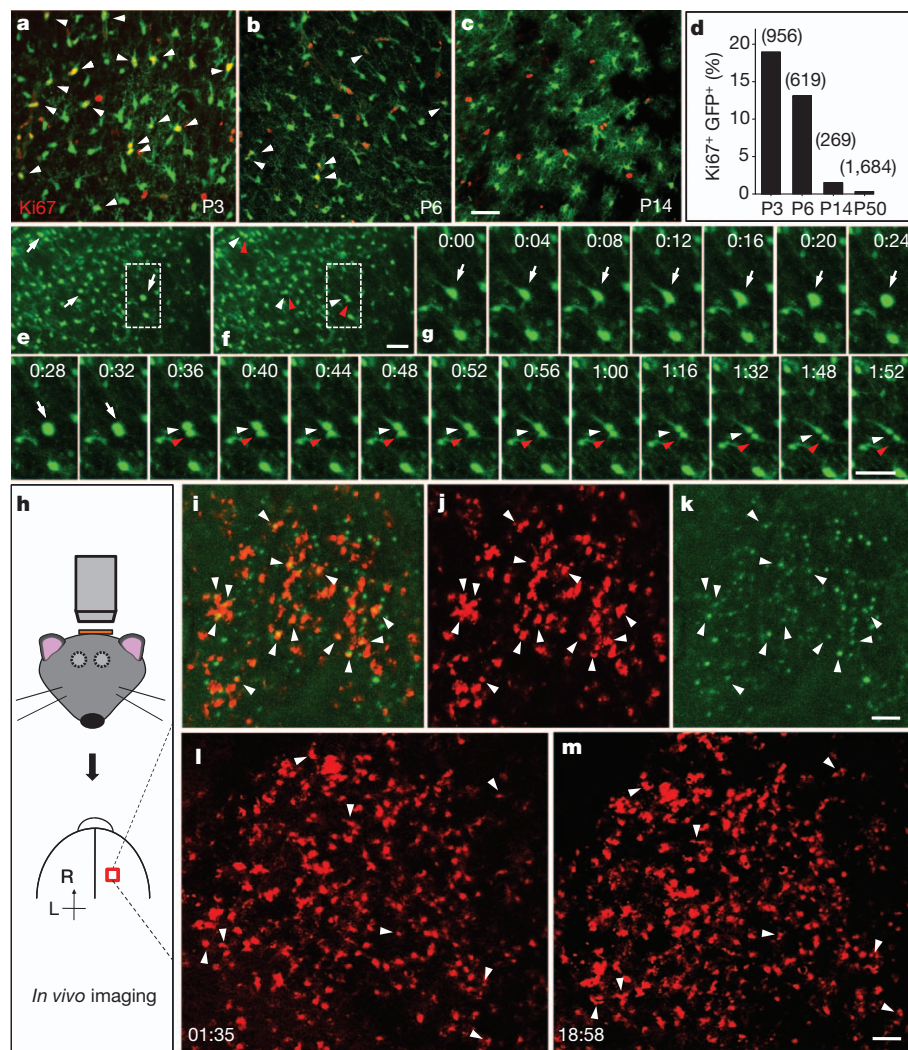


Figure 3 | Time-lapse imaging of local proliferation of astrocytes. a–c, Proliferating astrocytes (arrowheads) in the cortex of hGFAP-GFP transgenic mice, at P3 (a), P6 (b) and P14 (c). d, Summarized data for the percentage of Ki67⁺GFP⁺ cells among GFP⁺ cells with strong GFP signals. e, f, Sequential images of a cortical slice from a P3 hGFAP-GFP transgenic mouse (e, parent cells (arrows); f, daughter cells (arrowheads)). g, Time-lapse images

(1 h 52 min) of a dividing GFP⁺ cell in e and f. h, Procedure to image cell division *in vivo*. i–k, Images from a P4 triply transgenic hGFAP-CreER;Ai14;CAG-Fucci-Green mouse (i, combined images; j, tdTomato; k, mAG signal; arrowheads, dividing astrocytes). l, m, Time-lapse images at 1 h 35 min (l) and 18 h 58 min (m) from a P5 hGFAP-CreER;Ai14 transgenic mouse (arrowheads, dividing astrocytes). Scale bars, 40 μm (a–g) and 100 μm (k, m).

transgenic mice showing that NG2 glia generated very few astrocytes in the cortex during postnatal life (data not shown), our results reveal that Astro-like-D cells are the parent cells of locally generated astrocytes.

To assess the abundance of proliferating astrocytes within the cortex, we perfused hGFAP-GFP transgenic mice for Ki67 immunostaining and observed numerous astrocytes in the process of cell division before P10 (18.9% at P3, $n = 956$ GFP⁺ cells; 13.1% at P6, $n = 619$ GFP⁺ cells; 1.5% at P14, $n = 269$ GFP⁺ cells; 0.30% at P48–52, $n = 1,684$ GFP⁺ cells; Fig. 3a–d and Supplementary Fig. 9). To directly monitor local generation of astrocytes, we performed time-lapse imaging of acute cortical slices from hGFAP-GFP transgenic mice and found that roughly 2% of astrocytes divided within 3 h ($2.0 \pm 0.2\%$, 16 of 809 cells with strong GFP signals from five mice, P3–5; Fig. 3e–g and Supplementary Movie 1). Because NG2 glia maintain proliferative ability throughout life¹⁴, and some hippocampal NG2 glia are reported to have a very weak GFP signal in another hGFAP-GFP transgenic line²², we tested whether NG2 glia could have been among the dividing cells with a strong GFP signal, by loading biocytin into cells with a strong GFP signal through a recording pipette. We found that biocytin diffused only among cells with strong GFP signals ($n = 5$ slices; Supplementary Fig. 10), which is consistent with our observations from doubly transgenic hGFAP-GFP;NG2BacDsRed mice (Supplementary Fig. 11). Thus, the dividing cells with strong GFP expression are astrocytes rather than NG2 glia.

The density of dividing cells in acute brain slices showed no obvious change within a 6-h period after slice preparation (Supplementary Fig. 12); however, mature astrocytes could conceivably be induced to undergo local gliogenesis by means of a stab wound *in vivo*²³. We therefore performed *in vivo* imaging with an open skull but an intact pial surface within 1 h after surgery on the triply transgenic hGFAP-CreER;Ai14;CAG-Fucci-Green mice (Fig. 3h). We observed abundant dividing astrocytes (12.4%, tdTomato⁺mAG⁺; Fig. 3i–k, P3–6), a similar observation to that in brain sections (Fig. 3a, b, d). Because the thinned skull preparation does not cause astrocytic gliosis or activation of microglia²⁴, we then performed long-term time-lapse imaging with the thinned skull preparation in hGFAP-CreER;Ai14 transgenic mice and obtained similar results: abundant astrocytes were generated locally within the cortex (about 8% of progeny, $(8 \times 2)/212$ cells in Fig. 3l, m). Because of the difficulty of identifying a dividing cell under thinned skull if its two daughter cells did not separate completely, we probably underestimated the percentage of astrocytes produced on the basis of *in vivo* time-lapse imaging.

To determine whether dividing astrocytes in the cortex undergo symmetric division to produce astrocytes, or asymmetric division to generate multiple cell types as SVZ cells do^{8,25}, we recorded from their progeny during and shortly after cytokinesis (Fig. 4a, b). The two daughter cells shared similar *I*–*V* relationships that were characteristic of astrocytes (Fig. 4c). In addition, we examined the daughter-cell morphology in P6–8 hGFAP-CreER;Ai14 transgenic mice and found the two daughter cells occupying comparable areas and showing similar labelling with the astrocyte marker, BLBP (Fig. 4d and Supplementary Table 1). Thus, locally dividing astrocytes in the cortex primarily undergo symmetrical division to generate two daughter astrocytes.

To determine whether the progeny maintained their astrocytic fate after exiting from the cell cycle, we administered tamoxifen at P0–2 to hGFAP-CreER;Ai14 transgenic mice to label astrocytes permanently with tdTomato and examined their locally generated progeny 1 week later ($n = 4$ mice) to test whether these tdTomato⁺ cells still expressed the astrocyte marker BLBP. Although there were many progenitors or neurons with tdTomato expression in SVZ and hippocampal dentate gyrus (Supplementary Fig. 7f, g), nearly all of the tdTomato⁺ cells were BLBP⁺ or GFAP⁺ in the cortex (99.8%, motor and barrel cortex; Supplementary Fig. 7a–e). Because we found that few astrocytes would enter programmed cell death in the cortex (Supplementary Fig. 13), it is most likely that the progeny arising from local astrocyte division retained astrocytic identity long after exiting from the cell cycle. For

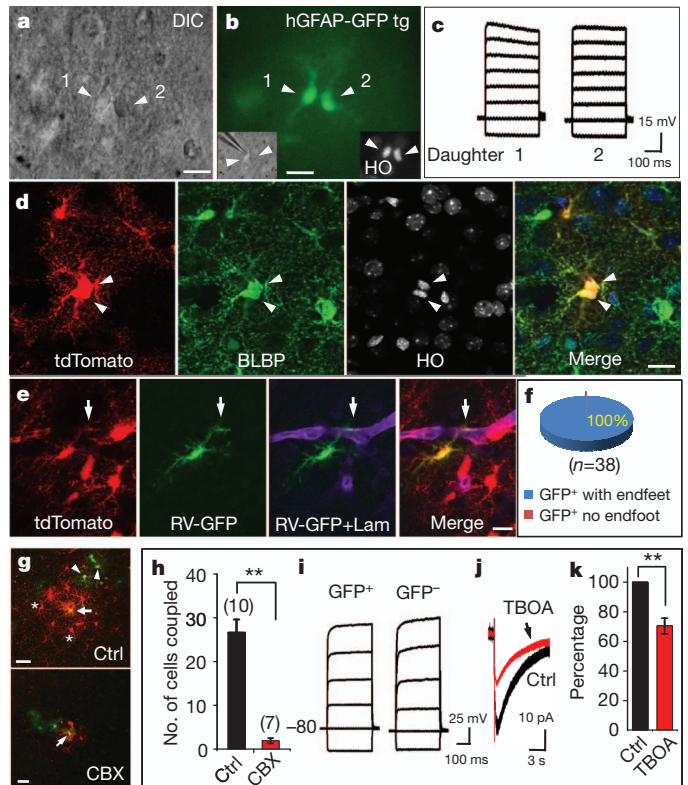


Figure 4 | Symmetric division of proliferating astrocytes and the function of their progeny. **a**, A pair of daughter astrocytes (arrowheads) at late telophase under differential interference contrast. **b**, Both cells had GFP signal. Nuclei were stained with Hoechst 33342 (HO, inset). **c**, Voltage responses of two daughter cells evoked by step currents (–1 to 6 nA). **d**, Two daughter cells in telophase (arrowheads) from a P8 hGFAP-CreER;Ai14 transgenic mouse. **e**, An astrocyte infected by GFP-expressing retroviruses (green) and expressing tdTomato (red) formed endfeet (arrows) with blood vessels (Laminin⁺, purple) in a P19 hGFAP-CreER;Ai14 transgenic mouse. Tamoxifen was injected at P2, and cells were infected with retrovirus at P5. **f**, The percentage of progeny cells marked by retroviruses (GFP⁺tdTomato⁺) that had endfeet (GFP⁺tdTomato⁺ with endfeet, blue). **g**, A retrovirus-infected astrocyte progeny (GFP⁺, green, arrow) in the absence (Ctrl, upper) or presence (lower) of 100 μ M carbenoxolene (CBX) was injected with biocytin (red). Without CBX, both GFP⁺ astrocytes (arrowheads) and GFP⁺ astrocytes (asterisks) contained biocytin (red), as a result of gap-junction coupling with the astrocyte progeny injected with biocytin. **h**, The number of cells coupled. Two asterisks, $P < 0.01$, (unpaired *t*-test). **i**, Current responses of uninfected (GFP[–]) and infected (GFP⁺) astrocyte progeny. **j**, **k**, Glutamate transporter current (**j**) and its summarized data (**k**) from infected astrocyte progeny before (black) and after (red) application of blocker TBOA (100 μ M, $70.4 \pm 5.3\%$, $n = 7$). Two asterisks, $P < 0.01$ (paired *t*-test). Scale bars, 10 μ m (**a**, **b**, **d**, **e**) and 20 μ m (**g**). Error bars indicate s.e.m.

further confirmation, tamoxifen was administered at P0–2 and retroviruses were injected locally at P3–5 (3 days after tamoxifen) in the cortex of hGFAP-CreER;Ai14 transgenic mice. The fate of doubly labelled cells (tdTomato⁺GFP⁺) was then assessed 2 weeks later. Because tdTomato marked cells that had expressed astrocyte markers before retroviral infection (Supplementary Fig. 7), the tdTomato⁺GFP⁺ cells correspond to the progeny of astrocytes that were infected by retroviruses during cell division. We found that all doubly labelled cells formed endfoot-like structures with blood vessels (38 of 38 yellow cells; Fig. 4e, f), a characteristic of differentiated astrocytes. These results demonstrate that locally generated progeny retain astrocytic identity long after they exit from the cell cycle.

We then asked whether daughter astrocytes arising from local astrocyte division integrate functionally into the existing glial network as mature astrocytes. The intercellular communication by means of gap junctions between astrocytes has a critical function in ion buffering

in the brain²⁶. At 1–3 weeks after viral infection, we loaded biocytin into infected astrocytes and found that biocytin diffused into neighbouring GFP⁺ astrocytes as well as some GFP⁺ astrocytes within 20 min (26.7 ± 2.9 cells, $n = 10$ slices; Fig. 4g–i and Supplementary Fig. 14). The coupling was inhibited by carbonoxelene (CBX), a blocker for gap junctions (1.9 ± 0.6 cells, $n = 7$ slices; Fig. 4g, h and Supplementary Fig. 14), indicating that locally generated astrocytes successfully integrated into existing glial networks. Another classical function of astrocytes is to clear glutamate from the synaptic cleft²⁷. After stimulation of neuronal fibres, an inward current with slow decay time course appeared in all the GFP⁺ astrocytes recorded (20 of 20 GFP⁺ cells, P12–19; Fig. 4j) and was sensitive to the glutamate transporter blocker TBOA (Fig. 4j, k). The remaining currents, lasting for more than 10 s (decay time 13.3 ± 0.4 s, $n = 4$; Fig. 4j), correspond to K_{ir} activation after neuronal excitation²⁸. These observations thus reveal that locally generated astrocytes function as mature astrocytes to take up glutamate and K⁺ ions after neuronal activity.

We have demonstrated that local generation of astrocytes within the postnatal cortex provides a major glial source, at least in layers I–IV, whereas astrocytes generated early in development are derived from radial glia^{4–6}, SVZ progenitors (including SVZ glioblasts and migratory glioblasts)^{4–6,20}. Once a subset of early astrocytes from those sources colonize and differentiate in the cortex as ‘pioneers’, local division of these differentiated astrocytes has a pivotal role in glial production after birth in the cortex.

Astrocytic endfeet almost fully cover the blood vessels by postnatal day 20 and are crucial to the regulation of cerebral blood flow²⁹ and the transport of nutrients from blood to neurons³⁰. It is not yet clear how this large number of locally generated astrocytes can coordinate with angiogenesis to form the complete gliovascular interface. Furthermore, aberrant gene activity affecting glial proliferation is one potential cause of gliomas, which comprise nearly 80% of primary malignant brain tumours³. It will also be of interest to test whether gliomas could have arisen from defective regulation of locally dividing glial cells in the brain.

METHODS SUMMARY

For live nuclear labelling, slices were incubated with Hoechst 33342 ($2 \mu\text{g ml}^{-1}$) as described previously¹⁴. For *in vivo* imaging, the pups with thinned skull (Fig. 3h, i, m) were immobilized with 4% agarose. The mouth of the pup was attached to a 1-ml pipette tip that was connected to a tube for inhalation. All data are given as means \pm s.e.m.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 22 August 2011; accepted 14 February 2012.

Published online 28 March 2012.

1. Azevedo, F. A. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532–541 (2009).
2. Eroglu, C. & Barres, B. A. Regulation of synaptic connectivity by glia. *Nature* **468**, 223–231 (2010).
3. Schwartzbaum, J. A., Fisher, J. L., Aldape, K. D. & Wrensch, M. Epidemiology and molecular pathology of glioma. *Nat. Clin. Pract. Neurol.* **2**, 494–503 (2006).
4. Cameron, R. S. & Rakic, P. Glial cell lineage in the cerebral cortex: a review and synthesis. *Glia* **4**, 124–137 (1991).
5. Marshall, C. A., Suzuki, S. O. & Goldman, J. E. Gliogenic and neurogenic progenitors of the subventricular zone: who are they, where did they come from, and where are they going? *Glia* **43**, 52–61 (2003).
6. Kriegstein, A. & Alvarez-Buylla, A. The glial nature of embryonic and adult neural stem cells. *Annu. Rev. Neurosci.* **32**, 149–184 (2009).
7. Bandeira, F., Lent, R. & Herculano-Houzel, S. Changing numbers of neuronal and non-neuronal cells underlie postnatal brain growth in the rat. *Proc. Natl Acad. Sci. USA* **106**, 14108–14113 (2009).
8. Levison, S. W. & Goldman, J. E. Both oligodendrocytes and astrocytes develop from progenitors in the subventricular zone of postnatal rat forebrain. *Neuron* **10**, 201–212 (1993).
9. Zerlin, M., Levison, S. W. & Goldman, J. E. Early patterns of migration, morphogenesis, and intermediate filament expression of subventricular zone cells in the postnatal rat forebrain. *J. Neurosci.* **15**, 7238–7249 (1995).

10. Smart, I. & Leblond, C. P. Evidence for division and transformations of neuroglia cells in the mouse brain, as derived from radioautography after injection of thymidine-H³. *J. Comp. Neurol.* **116**, 349–367 (1961).
11. Suh, H. *et al.* *In vivo* fate analysis reveals the multipotent and self-renewal capacities of Sox2⁺ neural stem cells in the adult hippocampus. *Cell Stem Cell* **1**, 515–528 (2007).
12. Guo, F., Ma, J., McCauley, E., Bannerman, P. & Pleasure, D. Early postnatal proteolipid promoter-expressing progenitors produce multilineage cells *in vivo*. *J. Neurosci.* **29**, 7256–7270 (2009).
13. Sakaue-Sawano, A. *et al.* Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487–498 (2008).
14. Ge, W. P., Zhou, W., Luo, Q., Jan, L. Y. & Jan, Y. N. Dividing glial cells maintain differentiated properties including complex morphology and functional synapses. *Proc. Natl Acad. Sci. USA* **106**, 328–333 (2009).
15. Ge, W. P. *et al.* Long-term potentiation of neuron–glia synapses mediated by Ca²⁺-permeable AMPA receptors. *Science* **312**, 1533–1537 (2006).
16. Wang, D. D., Krueger, D. D. & Bordey, A. Biophysical properties and ionic signature of neuronal progenitors of the postnatal subventricular zone *in situ*. *J. Neurophysiol.* **90**, 2291–2302 (2003).
17. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
18. Casper, K. B., Jones, K. & McCarthy, K. D. Characterization of astrocyte-specific conditional knockouts. *Genesis* **45**, 292–299 (2007).
19. Zhuo, L. *et al.* Live astrocytes visualized by green fluorescent protein in transgenic mice. *Dev. Biol.* **187**, 36–42 (1997).
20. Burns, K. A., Murphy, B., Danzer, S. C. & Kuan, C. Y. Developmental and post-injury cortical gliogenesis: a genetic fate-mapping study with Nestin-CreER mice. *Glia* **57**, 1115–1129 (2009).
21. Zhu, X. *et al.* Age-dependent fate and lineage restriction of single NG2 cells. *Development* **138**, 745–753 (2011).
22. Matthias, K. *et al.* Segregated expression of AMPA-type glutamate receptors and glutamate transporters defines distinct astrocyte populations in the mouse hippocampus. *J. Neurosci.* **23**, 1750–1758 (2003).
23. Buffo, A. *et al.* Origin and progeny of reactive gliosis: a source of multipotent cells in the injured brain. *Proc. Natl Acad. Sci. USA* **105**, 3581–3586 (2008).
24. Xu, H. T., Pan, F., Yang, G. & Gan, W. B. Choice of cranial window type for *in vivo* imaging affects dendritic spine turnover in the cortex. *Nature Neurosci.* **10**, 549–551 (2007).
25. Doetsch, F., Caille, L., Lim, D. A., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* **97**, 703–716 (1999).
26. Gutnick, M. J., Connors, B. W. & Ransom, B. R. Dye-coupling between glial cells in the guinea pig neocortical slice. *Brain Res.* **213**, 486–492 (1981).
27. Bergles, D. E. & Jahr, C. E. Synaptic activation of glutamate transporters in hippocampal astrocytes. *Neuron* **19**, 1297–1308 (1997).
28. Ge, W. P. & Duan, S. Persistent enhancement of neuron–glia signaling mediated by increased extracellular K⁺ accompanying long-term synaptic potentiation. *J. Neurophysiol.* **97**, 2564–2569 (2007).
29. Takano, T. *et al.* Astrocyte-mediated control of cerebral blood flow. *Nature Neurosci.* **9**, 260–267 (2006).
30. Rouach, N., Koulakoff, A., Abudara, V., Willecke, K. & Giaume, C. Astroglial metabolic networks sustain hippocampal synaptic transmission. *Science* **322**, 1551–1555 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. D. McCarthy and H. Zeng for providing us with the hGFAP-CreER and Ai14 transgenic mice, respectively; A. Sakaue-Sawano for the CAG-Fucci-Green transgenic mice; M. Stryker, Y. Xiang, X.-q. Wang, S. Barbel, J.-d. Chen, W. Zhou, F. Huang, and members of the Jan laboratory for discussion; G.-n. Li for help on electroporation; Y. Li for advice on retroviral experiments; and E. Unger, C. Guo, H. Yang, Q. Deng, J. Berg and X.-y. Li for reading the manuscript. W.-P.G. is a recipient of a Long-Term Fellowship of Human Frontier Science Program (HFSP) and National Institute of Neurological Disorders and Stroke (NINDS) Pathway to Independence Award. This work was supported by a NINDS K99/R00 award (1K99NS073735) to W.-P.G., a National Institute of Mental Health R37 grant (4R37MH065334) to L.Y.J., a National Institutes of Health (NIH) R01 grant (5R01MH084234) to Y.N.J., and grants from the NIH/National Institute on Aging P01 AG010435, MH090258, Jeffrey M. and Barbara Picower Foundation (JBP) and McDonnell Foundation to F.H.G. L.Y.J. and Y.N.J. are Howard Hughes Medical Institute investigators.

Author Contributions W.-P.G. conceived the project, designed and performed the experiments and analysed the data. L.Y.J. and Y.N.J. supervised the work and helped to design the experiments. W.-P.G. and L.Y.J. wrote the manuscript. F.H.G. and A.M. provided MLV retrovirus and CAG-Fucci-Green transgenic mice, respectively. All authors reviewed and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.Y.J. (lily.jan@ucsf.edu).

METHODS

Animals and tamoxifen administration. The CAG-Fucci-Green transgenic line was from A.M.'s laboratory, the hGFAP-CreER line was from K. D. McCarthy's laboratory (UNC), and the NG2BacDsRed transgenic line was from A. Nishiyama's laboratory. Both NG2-CreBac³¹ and NG2-CreER were generated in Nishiyama's laboratory and bought from Jackson Laboratory. The Ai14 transgenic mice were from H. Zeng's laboratory. Tamoxifen inductions were as described¹⁸. For induction in hGFAP-CreER;Ai14 transgenic mice, an intraperitoneal or subcutaneous injection of tamoxifen (dissolved in a 1:10 mixture of ethanol and sunflower oil) at 3 mg per 40 g of body weight was administered once at the time indicated. All animals were treated in accordance with protocols approved by the Institutional Animal Care and Use Committee at UCSF.

In vivo electroporation. Newborn to 2-day-old pups (P0–2) were anaesthetized by hypothermia (about 4 min) and fixed to a support with a sticking plaster. GFP complementary DNAs were cloned into the chicken β -actin CMV promoter-driven expression vector pCAGGS. DNA solution (1–2 μ l) prepared at 2 mg ml⁻¹ in 10 mM Tris-HCl pH 8.0, with 0.04% trypan blue, was injected into the lateral ventricle with a pulled-out glass capillary (diameter 50–100 μ m)³². Animals were subjected to five electric stimuli of 50 V, each lasting 50 ms, at 950-ms intervals using a square-pulse electroporator BTX830.

Retroviral preparation and in vivo infection. pCAG-GFP-PRE contains replication-defective murine leukaemia virus (MLV)-based retroviral elements designed to carry and express enhanced GFP under CMV promoter and CAG promoter (modified chicken β -actin promoter with enhanced sequences from CMV) with control of the MLV long terminal repeat. We followed the detailed protocol for preparation from Gage's laboratory³³. In brief, three plasmids (pCAG-GFP-PRE, pCMV-gp and CMV-vsvg) were transfected to HEK 293T cells with Lipofectamine 2000. Viruses containing supernatants were harvested 2 days after transfection by centrifugation twice at 65,000g for 2 h (Discovery 90SE; Sorvall). Final virus titres were about 10⁶–10⁷ colony-forming units ml⁻¹ as measured by infecting HEK 293T cells. Viruses with the GFP reporter gene were injected (1 μ l) into either C57BL/6 wild-type or hGFAPCreER;Ai14 transgenic mice at P0–9. For *in vivo* infection, pups were anaesthetized with ice for 3–5 min, and the injection was performed as described³⁴. After injection, the pups were put back in a cage with a lamp to keep them warm. They were returned to their home cage when fully recovered.

Immunocytochemistry. Mice were perfused with 4% paraformaldehyde in 1 \times PBS. Brains were cut into sections 25–50 μ m thick with a cryostat (model CM3050S; Leica). Floating sections were permeabilized with 0.25% Triton X-100 in 1 \times PBS and then blocked for 2 h with 5% BSA and 3% normal goat serum with 0.25% Triton X-100 in 1 \times PBS. Primary antibodies for Ki67 (1:200 dilution, rabbit, monoclonal; Thermo Scientific), BLBP (1:1,000, rabbit, polyclonal; Invitrogen) or Laminin (1:500, rabbit, polyclonal; Sigma) were applied to sections alone or in combination and left to incubate for 24–48 h at 4 °C. Together with DAPI or Hoechst 33342 (1 μ g ml⁻¹; Invitrogen), secondary antibodies conjugated with Alexa488, 555, 568 or 633 (1:750) were applied for 2 h at room temperature (22–25 °C). To identify apoptotic astrocytes, sections were incubated for 15 min with 1 μ g ml⁻¹ propidium iodide after the treatment with 0.2 mg ml⁻¹ RNase (DNase-free) in 1 \times PBS for 30 min at 37 °C as described previously^{35,36}.

Slice preparation. Slices were prepared as described previously¹⁵. In brief, after decapitation, mouse brains were dissected rapidly and sliced with a vibratome (VT-1000S; Leica) in ice-cold oxygenated (95% O₂ and 5% CO₂) artificial cerebrospinal fluid solution (aCSF) containing (in mM): 119 NaCl, 2.5 KCl, 2.5 CaCl₂, 1.3 MgSO₄, 1 NaH₂PO₄, 26.2 NaHCO₃ and 11 glucose. Transverse slices (250 μ m in thickness) were then maintained in an incubation chamber for at least 1 h at room temperature before whole-cell recording, nuclear dye loading or time-lapse imaging.

Electrophysiology and live cell nuclear labelling. Whole-cell recordings from mouse brain slices were conducted with the aid of markers (GFP or Hoechst 33342) to identify infected cells or dividing cells. Astrocytes in hGFAP-GFP transgenic mice were identified by bright green fluorescence under the

microscope. For live nuclear labelling, slices were incubated with Hoechst 33342 (diluted to 2 μ g ml⁻¹ in aCSF) at room temperature for 30 min as described previously¹⁴. Recording pipettes were routinely filled with a solution containing (in mM): 125 potassium gluconate, 15 KCl, 10 HEPES, 3 MgATP, 0.3 Na-GTP, 5 Na-phosphocreatine and 0.2 EGTA (pH 7.2–7.4, 290–300 mosM). For glutamate transporter currents, pipette solution contained (in mM): 125 caesium gluconate, 5 CsCl, 10 HEPES, 3 MgATP, 0.3 Na-GTP, 0.2 EGTA and 5 Na-phosphocreatine (pH 7.2–7.4, 290–300 mosM). Membrane potential in voltage-clamp mode was held at -80 mV. Current pulses (20–60 μ A, 0.1 ms, 0.05 Hz) were delivered through extracellular bipolar electrodes placed roughly 200–300 μ m from the cells being recorded to induce transporter current.

Biocytin labelling. Glial cells were filled with 0.1% biocytin (ϵ^N -biotinyl-L-lysine; Vector Lab) by means of a whole-cell recording electrode, as reported previously^{15,28}. Biocytin was dissolved in the recording pipette solution. Slices were fixed overnight with 4% paraformaldehyde at 4 °C before treatment for 2 h with blocking solution containing 5% BSA, 3% normal goat serum and 0.25% Triton X-100. Slices were then stained for 2 h with DyLight 549-conjugated streptavidin (1:1,000; Vector Lab). In Fig. 2o, DyLight 549 was added together with Alexa 633 (second antibodies against anti-Ki67) after washing out excess primary antibody against Ki67.

Confocal time-lapse imaging of acute brain slices. GFP⁺ cells at cortical slices from hGFAP-GFP transgenic mice (P3–5) were imaged on a Zeiss LSM510 two-photon confocal microscope equipped with objective 20 \times /0.5W and 63 \times /0.9W (Zeiss). Cells were scanned with xyz mode (four optical slices in z, with 8- μ m interval between slices). The frame interval was 4 min for 30–100 frames. Projection images were made from z-stacks that included all individual GFP⁺ cells. During imaging, slices were kept in a chamber with perfusion of aCSF (see above) at 32–34 °C.

Confocal time-lapse imaging in vivo. The pups (P3–6, hGFAP-CreER;Ai14 transgenic pups; Fig. 3l, m) were anaesthetized by hypothermia: 4–5 min in ice wrapped in a piece of cloth. A small fraction of skin (3 mm \times 3 mm) was removed over an area to be imaged. The pups then were returned to a box for 3–4 h until the incision site healed (no bleeding). A high-speed micro-drill was used to thin a circular area of skull, typically about 1 mm in diameter. The mouth of the pup was attached to a 1-ml pipette tip that was connected to a tube for inhalation. Pups were then immobilized with 4% agarose. Imaging was performed using a two-photon laser-scanning microscope based on a mode-locked laser system operating at 930 nm, equipped with one of the following objectives: 10 \times , 0.25 numerical aperture (NA); 20 \times , 0.8 NA collected emission more than 560 nm for tdTomato and 500–550 nm for mAG. Sometimes tdTomato was excited with a laser at 543 nm. Images were taken every 1.5 h for the first 3 h, and then the pups were put back in a box and allowed to move freely. Additional images were taken every 9–12 h for the following 18–24 h. During imaging, pups were fully anaesthetized with 2–4% isoflurane for 4–5 min. Cells under the pia were scanned with xyz mode (16 slices in z with 10- μ m interval between optical slices.). Shortly after an imaging session, isoflurane was turned off and oxygen was left on until the animal fully recovered. For Fucci-Green;hGFAP-CreER;Ai14 pups (Fig. 3i–k) we removed a small fraction of the skull (1 mm²) and images were taken within 1 h after surgery.

31. Zhu, X., Bergles, D. E. & Nishiyama, A. NG2 cells generate both oligodendrocytes and gray matter astrocytes. *Development* **135**, 145–157 (2008).

32. Li, G. *et al.* Regional distribution of cortical interneurons and development of inhibitory tone are regulated by Cxcl12/Cxcr4 signaling. *J. Neurosci.* **28**, 1085–1098 (2008).

33. Tashiro, A., Zhao, C. & Gage, F. H. Retrovirus-mediated single-cell gene knockout technique in adult newborn neurons *in vivo*. *Nature Protocols* **1**, 3049–3055 (2006).

34. Merkle, F. T., Mirzadeh, Z. & Alvarez-Buylla, A. Mosaic organization of neural stem cells in the adult brain. *Science* **317**, 381–384 (2007).

35. Barres, B. A. *et al.* Cell death and control of cell survival in the oligodendrocyte lineage. *Cell* **70**, 31–46 (1992).

36. Krueger, B. K., Burne, J. F. & Raff, M. C. Evidence for large-scale astrocyte death in the developing cerebellum. *J. Neurosci.* **15**, 3366–3374 (1995).

A unique regulatory phase of DNA methylation in the early mammalian embryo

Zachary D. Smith^{1,2,3*}, Michelle M. Chan^{1,4*}, Tarjei S. Mikkelsen^{1,2}, Hongchang Gu¹, Andreas Gnirke¹, Aviv Regev^{1,5} & Alexander Meissner^{1,2,3}

DNA methylation is highly dynamic during mammalian embryogenesis. It is broadly accepted that the paternal genome is actively depleted of 5-methylcytosine at fertilization, followed by passive loss that reaches a minimum at the blastocyst stage. However, this model is based on limited data, and so far no base-resolution maps exist to support and refine it. Here we generate genome-scale DNA methylation maps in mouse gametes and from the zygote through post-implantation. We find that the oocyte already exhibits global hypomethylation, particularly at specific families of long interspersed element 1 and long terminal repeat retroelements, which are disparately methylated between gametes and have lower methylation values in the zygote than in sperm. Surprisingly, the oocyte contributes a unique set of differentially methylated regions (DMRs)—including many CpG island promoters—that are maintained in the early embryo but are lost upon specification and absent from somatic cells. In contrast, sperm-contributed DMRs are largely intergenic and become hypermethylated after the blastocyst stage. Our data provide a genome-scale, base-resolution timeline of DNA methylation in the pre-specified embryo, when this epigenetic modification is most dynamic, before returning to the canonical somatic pattern.

Cytosine methylation in mammals is an epigenetic modification that is largely restricted to CpG dinucleotides and serves multiple critical functions, including stable repression of target promoters, maintaining genomic integrity, establishing parent-specific imprinting patterns, and silencing endogenous retrotransposon activity^{1,2}. In somatic tissues, CpG methylation exhibits global patterns based on relative CpG density: CpG islands at housekeeping or developmental promoters are largely unmethylated, whereas non-regulatory CpGs distributed elsewhere in the genome are largely methylated^{1,3}. This landscape is relatively static across all somatic tissues, where most of the methylated CpGs are pre-established and inherited through cell division. Generally, only a small fraction of CpGs switch their methylation status as part of an orchestrated regulatory event^{4–7}.

DNA methylation is much more dynamic during mouse germ cell and pre-implantation development. The classical model postulates that at fertilization, a targeted, although widespread, catalytic process actively removes DNA methylation contributed by the paternal gamete. Recent evidence implicates a demethylation mechanism that transitions through a hydroxymethylated intermediate that is catalysed by the Tet3 member of the Tet family^{8,9}. However, only a portion of hydroxylated targets seems to be actively catalysed to complete demethylation, and the identity of these targets remains unknown¹⁰. After fertilization, there appears to be a passive loss of global DNA methylation levels that continues until the blastocyst stage, where the inner cell mass (ICM) that gives rise to the embryo proper is first specified (reviewed by ref. 11). Recent evidence indicates that this passive depletion may also be facilitated in part by Tet-enzyme-mediated hydroxylation¹⁰. After specification of the ICM, the embryo implants into the uterine lining in concert with gastrulation, which

co-occurs with global remethylation of the genome that is believed to contribute to lineage restriction and the loss of cellular potency^{12,13}.

Unfortunately, on a quantitative, genome-wide scale, little is known about the specific dynamics of cytosine methylation during these earliest developmental stages¹⁴. The classical model is drawn from observations made using either global measurements, such as immunohistochemistry, or from limited analysis at individual loci^{11,12,15–22}. Key questions about DNA methylation patterns in early development remain open, including which genomic features are specifically targeted, as well as the identities of DMRs inherited from either gamete beyond known imprint control regions (ICRs). Here we use genomic high-resolution methylation profiling^{23,24} to gain insight into the underlying mechanisms and regulatory principles of DNA methylation as it functions in early mammalian development.

Genome-scale methylation maps of murine embryogenesis

To generate a global and high-resolution view of early mammalian DNA methylation dynamics, we collected oocytes and sperm, as well as zygote, 2-, 4- and 8-cell cleavage stage embryos, the early ICM and embryonic day (E)6.5/7.5 post-implantation embryos (Fig. 1a and Supplementary Figs 1 and 2). All samples were extensively washed and purified to remove any somatic or gametic contaminants; maternal biasing from meiotic polar bodies (representing a 1× or 0.5× static genomic contaminant, respectively) was excluded by mechanical biopsy (Supplementary Fig. 1 and Supplementary Movie 1) and was further confirmed by assessing the paternal (129X1/SvJ) to maternal (C57BL/6 × DBA/2) ratio of known single nucleotide polymorphisms (SNPs) (Supplementary Figs 3 and 4). We

¹Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. ²Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA. ³Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ⁴Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁵Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

*These authors contributed equally to this work.

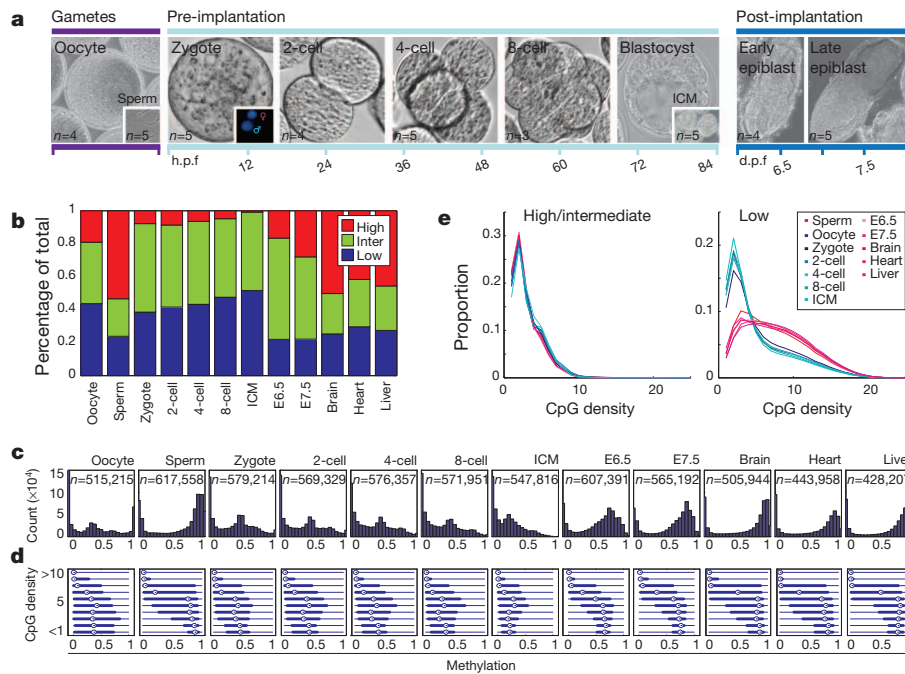


Figure 1 | Global CpG methylation dynamics across early murine embryogenesis. **a**, Samples isolated for methylation analysis with replicate number (n) highlighted. d.p.f., days post fertilization; h.p.f., hours post fertilization. **b**, Fraction of 100-bp tiles with high (≥ 0.8 , red), intermediate (inter, >0.2 and <0.8 , green) and low (≤ 0.2 , blue) methylation values. Brain, heart and liver tissue are shown for adult comparisons. **c**, Histogram of methylation values across 100-bp tiles. n is the number of tiles for each stage.

generated reduced representation bisulphite sequencing (RRBS)⁴ libraries from each stage to provide a comprehensive timeline of DNA methylation patterns during early mouse embryogenesis.

Compared to all other genome-wide profiling strategies currently available, RRBS is optimally suited for the low cell numbers that can be obtained from embryonic samples^{23–25}. Within our range of ~ 0.5 – 10 ng genomic DNA, RRBS provides the expected genomic coverage and high reproducibility (Supplementary Fig. 2). On average, we obtained the methylation status of 1,062,216 CpGs for comparative analysis (Supplementary Table 1). Bisulphite sequencing cannot distinguish between methyl- and hydroxymethylcytosine (hmC), and current methods for global profiling of hmC lack the sensitivity to investigate the pre-implantation stages in this study^{9,26–31}. Thus, we cannot draw any definitive conclusions regarding the base resolution hmC distribution, but this modification has not yet been linked to a regulatory mechanism other than to potentiate demethylation³². Given this ambiguity, regions of high mC/hmC methylation, especially those retained over multiple time points, could still be expected to function as if methylated.

The global embryonic pattern is unique

Current models postulate a phase of global hypomethylation during mammalian pre-implantation development that reaches a minimum at the morula/blastocyst stage. However, it is unknown which genomic regions are affected or how similar the embryonic methylation pattern is to the adult¹. To address these questions, we investigated the global dynamics of CpG methylation using 100-base-pair (bp) tiles (Methods). We found that oocytes are already globally hypomethylated compared to sperm (0.32 median methylation in oocyte versus 0.83 in sperm; Supplementary Fig. 5). We examined the relative proportion of genomic regions at each stage falling into high (≥ 0.8), intermediate (>0.2 and <0.8) or low (≤ 0.2) methylation categories. Notably, oocyte methylation levels more closely resemble those of early embryonic time points than the levels in sperm, post-implantation

d, Box plots of methylation values across local CpG densities highlight the difference between hypomethylated pre-implantation tissues and the somatic pattern seen in sperm, post-implantation and adult samples. Circle indicates the median, edges the 25th/75th percentile and whiskers the 2.5th/97.5th percentile. **e**, CpG density distribution for >0.2 methylation (left panel) and ≤ 0.2 methylation (right panel) tiles in stages that display somatic versus embryonic patterning (red and blue lines, respectively).

embryos, or adult tissues (Fig. 1b). We also observed a gradual increase in the fraction of tiles that exhibit intermediate and low methylation values from oocytes to the early ICM, which is consistent with loss of methylation over multiple cleavage divisions (Fig. 1b).

Sperm and post-implantation embryos show a strong inverse relationship between CpG density and methylation levels that is characteristic of somatic cells. In oocyte and pre-implantation samples, this dependence is weaker (Fig. 1c, d). In both pre- and post-implantation embryos, methylated CpGs (>0.2) tend to occur in low CpG density regions, as they do in somatic cells (Fig. 1e, left). However, the alternative relationship between higher CpG density and low methylation is not as apparent in the oocyte or the pre-implantation embryo (Fig. 1e, right). In summary, pre-implantation development represents a unique developmental period where methylation is differentially positioned and regulated before being restored and maintained in a somatic fashion.

Two transitions in early development

We next searched for substantial changes in regional DNA methylation through development. For each pair of consecutive stages, we compared methylation levels of each tile and classified it as changed if the difference exceeded 0.2 and was significant according to a false discovery rate (FDR)-corrected t -test. The most dramatic changes in DNA methylation occurred during two developmental transitions: between sperm and the zygote and between the early ICM and the post-implantation embryo (Fig. 2a). At each of these transitions, most changes were unidirectional (Fig. 2b): a gross reduction upon fertilization (mean = 0.47 decrease for 37% of tiles examined) and massive remethylation from the ICM onwards (mean = 0.46 increase in methylation at 66% of tiles). Within E6.5 and E7.5 post-implantation embryos, the methylation levels at most of the assayed tiles were stable or increased slightly (Fig. 2b). However, more subtle global changes, reflecting a gradual decrease in methylation, were observed from zygote/early cleavage through to the 8-cell stage and

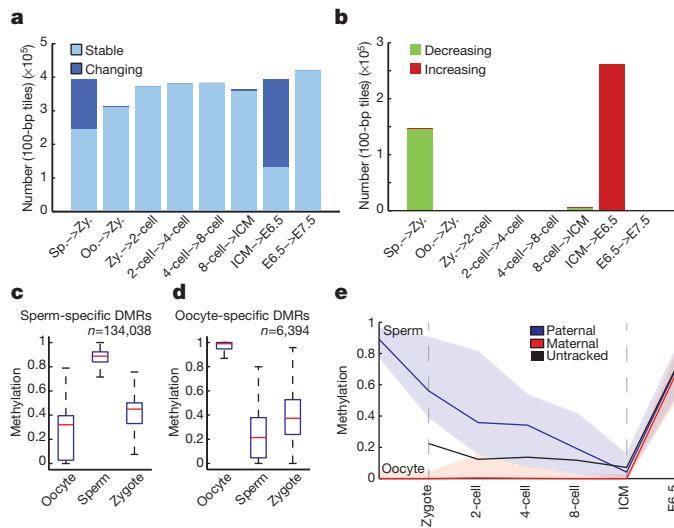


Figure 2 | Major transitions in DNA methylation levels during early development. **a**, 100-bp tiles available for pairwise comparison across consecutive embryonic stages. Tiles that remain unchanged (stable) at the indicated transitions are shown in light blue. Tiles that change by greater than 0.2 and are significant by *t*-test are highlighted in dark blue. **b**, 100-bp tiles with increasing (red) or decreasing (green) methylation levels at each consecutive transition show that major transitions are largely unidirectional. **c**, Box plot of methylation levels for sperm-specific DMRs ($n = 134,038$ tiles). Red line indicates the median, edges the 25th/75th percentile and whiskers the 2.5th/97.5th percentile. **d**, Box plot of methylation levels for oocyte-specific DMRs ($n = 6,394$ tiles) as in **c**. **e**, Seventy-four CpGs within sperm-specific DMR tiles (**c**) could be ascribed to paternal or maternal alleles and tracked across stages. Paternal CpG methylation values (blue line, median; coloured space, 25th/75th percentile) decrease by the zygote stage whereas maternal CpG methylation (red line, median; coloured space, 25th/75th percentile) remains unchanged. If untracked, these CpGs have an intermediate methylation value between those ascribed to a parent of origin (black line).

into the ICM, where methylation levels reached their lowest observed values (Fig. 1b, c).

The oocyte defines the early methylation landscape

Active demethylation is expected to occur before pronuclear fusion or the completion of DNA synthesis^{11,33}. When we compare methylation patterns between sperm and zygote, most regions in the genome show reduced methylation in the zygote with few sizeable changes in 2-cell embryos (Fig. 2b). Notably, the vast majority of tiles that are methylated at significantly different levels between gametes show higher methylation levels in sperm than in oocyte and many are reduced to levels near those of the oocyte (Fig. 2c, d). Using SNPs, we confirmed this observation by tracking 74 CpGs that fell within these tiles and could be assigned paternal- or maternal-specific values. Zygotes displayed a decrease in paternal methylation in contrast to maternally contributed CpGs, which remained unmethylated (Fig. 2e). Zygotes isolated here are likely in earlier stages of S phase, such that either a passive, replication-based mechanism could result in the synthesis of unmethylated, nascent DNA or DNA methylation could be removed by a targeted process^{10,34–36}. The similarities in methylation levels between zygote and the 2-cell stage, which represents one complete round of replication, is indicative that at least some observed demethylation is a consequence of targeted removal, but distinguishing between these two models may be complicated by the coupling of proposed base-excision repair mechanisms and DNA replication itself³⁵.

The few regions that are significantly hypermethylated in oocyte compared to sperm exhibit intermediate values in the zygote, suggesting a more direct inheritance of the allelic methylation state (Fig. 2d). The disparity in the zygotic resolution of regions that are differentially

methylated between the gametes indicates that the oocyte largely reflects the zygotic/pre-implantation methylome and prescribes its architecture (Supplementary Fig. 6). Thus, the oocyte methylome, rather than the sperm methylome, seems to be more reflective of patterns in the early embryo.

Retroelement dynamics at fertilization

Consistent with a demethylation model, we confirmed that the vast majority (96%) of tiles that are hypermethylated in sperm in our data set become less methylated in the zygote. Most of these tiles already exhibit lower methylation in the oocyte, such that additive effects could explain more subtle decreases in many regions. Interestingly, tiles exhibiting the most extreme methylation changes during the sperm to zygote transition are enriched for long interspersed elements (LINEs) ($P < 4.7 \times 10^{-184}$, $FDR < 0.05$, hypergeometric enrichment) (Fig. 3a and Supplementary Table 2). We directly estimated the methylation level for individual LINEs³⁷ surveyed by RRBS at each stage and found that changes in these elements are markedly bimodal during the sperm to zygote transition, with 18% of LINEs reducing their methylation values by over 0.45 (Fig. 3a). By comparison, 10% of captured long terminal repeat (LTR) retroelements exhibit similar levels of demethylation, but the distribution is not as clearly bimodal (Fig. 3b and Supplementary Table 3). Short interspersed elements (SINEs) are generally less methylated in sperm than other repeat classes³⁸, and also exhibit shifts in their methylation values from sperm to early embryo, but without the apparent bimodality observed for LINE elements (Supplementary Fig. 7).

Surprisingly, LINEs that changed most dramatically during the sperm to zygote transition largely consisted of two closely related families of L1 LINEs: L1Md_T and L1Md_Gf (Fig. 3c, d, $P < 4.7 \times 10^{-184}$, $P < 7.9 \times 10^{-6}$, hypergeometric enrichment test)^{37,39}. Repeats from these families had the largest and most consistent decrease, whereas those from other equally represented families, such as L1Md_A elements, showed smaller changes upon fertilization and maintained higher methylation values in both oocyte and zygote (Fig. 3e and Supplementary Fig. 8). Similarly, several LTR families showed considerable loss of methylation within the zygote (Fig. 3f, g), whereas the class II intracisternal A-particles (IAPs, Fig. 3h) did not. The latter finding is consistent with the known retention of high methylation levels of IAPs throughout cleavage^{19,22}.

Interestingly, during early development, all retrotransposons resolved identically, reaching minimal values at the ICM stage before increasing to more somatic levels by E6.5/7.5 (Fig. 3i). Thus, repeat elements exist in a less methylated state primarily in the oocyte pre-implantation stages (Supplementary Fig. 7). Bisulphite sequencing cannot address if methylated cytosines at these repeats are converted to hmCs before a subset is further targeted for complete demethylation. Some mCs may be targeted for active demethylation via this intermediate form, whereas the remaining mC/hmC residues may lose their methylation more passively through cleavage, consistent with recent metaphase immunostaining results¹⁰.

Gametes confer distinct features as DMRs

Although loss of methylation is widespread, some epigenetic information must be differentially contributed by the two gametes, including known ICRs that maintain their allele-specific methylation pattern throughout embryogenesis⁴⁰. We systematically searched for inherited DMRs contributed from either gamete by applying linear regression to all tiles that had mean methylation ≥ 0.75 in one gamete and ≤ 0.25 in the other. We identified 376 oocyte-contributed DMRs with intermediate methylation levels in the zygote ($P < 0.047$, $FDR < 0.05$, ANOVA; linear regression residual < 0.29 , $FDR < 0.1$; Fig. 4a) and 4,894 sperm-contributed DMRs (Fig. 4c). Notably, oocyte-contributed DMRs reside primarily in CpG island-containing promoters (Fig. 4b and Supplementary Table 4), whereas sperm-contributed DMRs are predominantly intergenic (Fig. 4d). The sperm- and oocyte-contributed

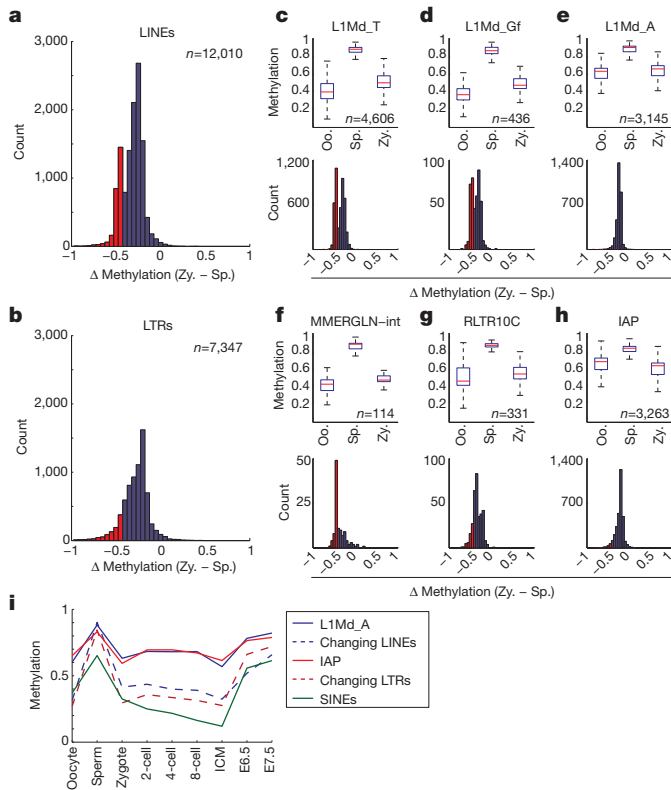


Figure 3 | Specific families of LINE and LTR retroelements exhibit the most dramatic methylation changes in the sperm to zygote transition.

a, Histogram of the difference in methylation levels, where negative values represent tiles decreasing from sperm to zygote, within LINE retroelement features that are captured by RRBS. 85% of the elements have a significant difference ($P < 0.04$, FDR < 0.05 ; t -test). The distribution is bimodal with 18% of elements displaying a change in methylation status ≥ 0.45 as highlighted in red. **b**, Differences in methylation between sperm and zygote within annotated LTR retroelements. Compared to LINEs, a smaller fraction of elements appear to be regulated by DNA demethylation (61% significant, 10% of those sampled exhibiting changes ≥ 0.45 as highlighted in red). **c–e**, Box plots of methylation levels in oocyte, sperm and zygote (top panels) as well as the distributions of change in methylation levels between sperm and zygote (bottom panels) for specific LINE-1 families, including those that are (c, d) or are not dynamic (e). Top panels: the red line indicates the median, edges the 25th/75th percentile and whiskers the 2.5th/97.5th percentile. Bottom panels: members of each family that are demethylated by greater than 0.45 are highlighted in red. **f–h**, Box plots of methylation levels in oocyte, sperm and zygote (top panels) and the distributions of change in methylation levels between sperm and zygote (bottom panels) for specific families of LTR retroelements, including MMERGLN (f), RLTR10C (g) and IAP elements (h). Top and bottom panels as in c–e. **i**, Mean methylation level for all elements of the L1Md_A LINE (solid blue line) and IAP LTR class (solid red line) that do not markedly change contrasted by LINEs (dashed blue line) and LTR elements (dashed red line) that show the greatest loss at fertilization. SINE elements (green line) are less methylated in sperm than other repeat elements and appear to decrease to oocyte levels.

ted DMRs also differed substantially in their relative CpG densities (Supplementary Fig. 9).

We next focused specifically on oocyte-contributed promoter DMRs, in part due to their unusual enrichment for high CpG-containing promoters (HCPs). Although this set had no clear functional enrichment, it did include the promoters of several interesting genes that are not expressed during later stages of oogenesis, such as *Dnmt3b* and the somatic isoform of *Dnmt1* (refs 41–43), which suggests a possible regulatory function for at least some of these DMRs. The use of genotyped strains allowed us to confirm that the methylation proximal to the CpG island promoter of copine VII (*Cpne7*), another putative DMR, was directly inherited from the oocyte

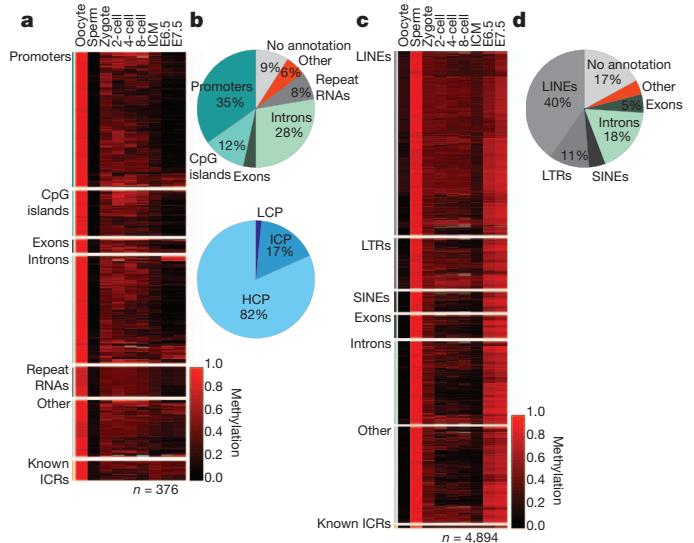


Figure 4 | Differentially methylated regions represent discrete gamete-specific feature classes. **a**, Heat map of methylation levels (black, 0; red, 1; grey, missing value) in 376 identified 100-bp tiles (rows) that behave as oocyte-contributed DMRs in the zygote. Tiles are sorted by functional classes and clustered within each class. Fifteen known ICRs, shown at the bottom, behave similarly in the early embryo and retain intermediate methylation through implantation. Other includes both Other and No annotation. **b**, Genomic features (top) and promoters of different CpG densities (bottom) in oocyte-contributed DMRs. Top: oocyte DMRs are enriched for promoters. Bottom: most of the 105 promoters that overlap oocyte-contributed DMR tiles are high CpG density promoters containing CpG islands (HCPs, light blue). **c**, Heat map of methylation levels (black, 0; red, 1; grey, missing value) in 4,894 identified 100-bp tiles (rows) that behave as sperm-contributed DMRs in pre-implantation embryos. Tiles are sorted and highlighted as in a. **d**, Genomic features in sperm-contributed DMRs are generally intergenic.

(Fig. 5a). As a set, oocyte-contributed promoter DMRs retained intermediate methylation values from the zygote through to the ICM, followed by resolution to hypomethylation in the specified embryo (Fig. 5b, c). Thus, CpG island methylation is transiently stabilized during cleavage divisions before re-establishing an unmethylated state around implantation. A distinct methylation pattern during pre-implantation development is also observed in sperm-contributed DMRs, which retain intermediate methylation values through to the ICM, before being hypermethylated post-implantation to typical somatic levels (Fig. 5d).

RRBS is designed to enrich for CpG dinucleotides (sixfold), but it also captures the other three non-CpG dinucleotides at normal frequencies. Of these, CpA is the predominant target for methylation in mouse and human^{44,45}. Consistent with previous locus-specific observations^{46,47}, we found that oocytes had the highest global CpA methylation level observed across pre-implantation development, and that this level decreased by $\sim 50\%$ in the zygote stage (Fig. 5b and Supplementary Fig. 10). This indicates that non-CpG methylation is inherited as part of oocyte-contributed methylated alleles but is likely lost more rapidly.

Discussion

To better understand the regulation of methylation patterns during its most dynamic phase, we generated genome-scale maps of DNA methylation in both gametes and through implantation. We find that methylation contributed by sperm to the zygote is altered most within retroelements of specific families, whereas other elements remain more protected and retain higher methylation levels throughout development (Supplementary Fig. 11). The methylation status of the oocyte is a strong predictor of levels in the zygote, and regions that are already hypomethylated in the oocyte could explain much of the disparity between the early embryo and sperm. The mechanism and targets of DNA demethylation

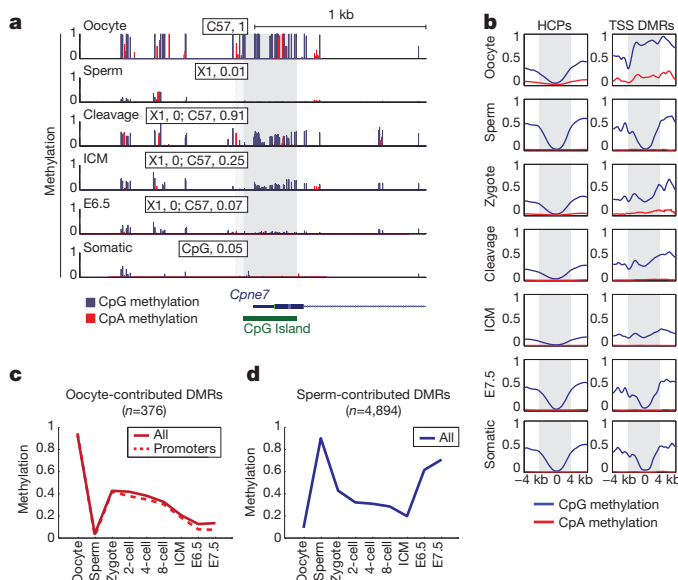


Figure 5 | DMRs resolve after cleavage to univalent hyper- or hypomethylated values in a gamete-of-origin-specific fashion. **a**, Single CpG resolution methylation within 2 kilobases (kb) of the *Cpne7* promoter in gametes and across embryonic development (rows). Dark grey bar highlights the CpG island. A CpG proximal to the island can be tracked to a phase resolving SNP and is highlighted in light grey, with paternal (X1) and maternal (C57) methylation values included as an inset for each trackable phase. Values for SNP methylation in 'cleavage' correspond exactly to those captured in the zygote. Blue bars, CpG methylation; red bars, CpA methylation. **b**, Composite plot of CpG (blue) and CpA (red) methylation for all HCPs (left) and for promoters that are specifically hypermethylated in oocytes (transcription start site (TSS) DMRs, right). The region ± 2 kb of the TSS is marked in grey. Identified promoter DMRs contributed by the oocyte are hypermethylated around the periphery of the TSS and resolve to intermediate values throughout cleavage. An expected HCP methylation architecture is re-acquired for these DMRs around implantation. **c**, Mean methylation levels across stages for oocyte-contributed DMRs in promoters (red, dashed line) versus the complete set (red, solid line). **d**, Sperm-contributed DMRs (blue line) generally resolve to hypomethylation.

during female gametogenesis could possibly be similar to those at work during fertilization³⁴. Regardless, the embryonic pattern more closely resembles that of the oocyte until the later stages of pre-implantation, where DNA methylation is further decreased.

In addition to classical ICRs, which exhibit parent-of-origin-specific methylation maintained through adulthood, a substantial number of CpG island promoters are specifically hypermethylated in the oocyte, in agreement with a recent study²⁵. Surprisingly, these regions retain intermediate values indicating differential allelic methylation before gradually decreasing through ICM specification and gastrulation, where somatic methylation patterns are re-established (Supplementary Fig. 11).

It remains to be investigated whether the diverse targets that exhibit low methylation levels during embryogenesis are the consequence of a single regulatory principle. LINE and LTR activity in the early embryo is associated with some of the earliest transcriptional events during zygotic genome activation. Targeted depletion by antisense oligonucleotides of the L1Md_T class as well as certain LTRs suggests a general requirement for retrotransposon transcription for progression through cleavage^{48,49}. These observations may also support data indicating that the elongation factor/histone acetyltransferase ELP3 is a component of the DNA demethylation machinery and could explain a tight relationship between complete demethylation and transcription-associated complexes⁵⁰.

It is likely that current interest in hmC will spur technical improvements that will permit quantitative dissection of mC and hmC patterns, which will help answer remaining questions regarding

Tet3's universal necessity for conversion to unmethylated cytosines, as well as the effect hmC may have on Dnmt-mediated inheritance³². Tet3's global conversion to hmC of the paternal genome does not seem to lead to equivalently marked demethylation on the basis of the retention of bisulphite-detected signal. The feature-specific dynamics of DNA methylation at fertilization indicate that Tet3 and hmC may be required for targeted demethylation, as well as for driving a gradual depletion through cleavage^{9,10}. Further experiments will be required to characterize this division-dependent demethylation in more detail, and expand it to regions with lower G+C content that are under-represented in RRBS. Notably, other mechanisms must retain heritable methylation information because many targets display relative epigenetic stability from zygote onward and some of these features exhibit embryogenesis-specific methylation patterns.

Our genome-scale single-base-resolution data provide improved understanding of the relationship between DNA methylation and early development. This expands earlier models derived from immunohistochemistry-based observations and begins to address remaining unanswered questions, setting the stage for future epigenetic studies in early mouse development.

METHODS SUMMARY

Gametes, cultured cleavage stage embryos, immunosurgically dissected E3.5 ICM, and post-implantation embryos were isolated as described previously (see Methods). Samples were purified through sequential KSOM microdrops (Millipore) and polar body contaminants mechanically dissected using XY laser (Hamilton Thorne) assisted biopsy (Supplementary Fig. 1 and Supplementary Movie 1). Reduced representation bisulphite libraries were generated from proteinase-K-purified, MspI-digested genomic DNA and sequenced using the Illumina Genome Analyzer II platform. Sequenced reads were aligned to the Mouse Genome Build 37 (mm9) using a custom computational pipeline that accounted for the strain identity of samples, which were either inbred or hybrid strains to provide adequate SNP tracking. Sampled cytosines covered $\geq 10\times$ were used for single CpG analysis. Alternatively, single CpGs were incorporated into features taken from ref. 4 or into 100-bp tiles using a $5\times$ threshold. Methylation levels reported for a sample is the average across replicates that met these threshold criteria. Tiles were considered to show a change between two stages if they exhibited a methylation difference ≥ 0.2 and statistical significance through a *t*-test after correction for multiple hypothesis testing (FDR < 0.05) using the Benjamini–Hochberg method. Retrotransposon annotations are from the RepeatMasker track of the UCSC genome browser. Novel DMRs were identified from a pool of 100-bp tiles where one gamete had a mean methylation ≥ 0.75 and the other had a mean methylation ≤ 0.25 . Linear regression applied to this set identified tiles that had zygotic methylation values that fell halfway between those of oocyte and sperm. SNPs between 129X1/Svj paternal and BDF1 (C57BL/6 \times DBA/2) maternal genomes were taken from Mouse Genome Informatics and used to assess relative maternal contamination as well as to identify the parent of origin in order to track allelic methylation values in DMRs and sites exhibiting demethylation.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 July 2011; accepted 15 February 2012.

Published online 28 March 2012.

- Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
- Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.* **33** (Suppl.), 245–254 (2003).
- Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nature Rev. Genet.* **9**, 465–476 (2008).
- Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
- Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature Genet.* **37**, 853–862 (2005).
- Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genet.* **39**, 457–466 (2007).
- Ji, H. *et al.* Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467**, 338–342 (2010).

8. Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Commun.* **2**, 241 (2011).
9. Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
10. Inoue, A. & Zhang, Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* **334**, 194 (2011).
11. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093 (2001).
12. Kafri, T. *et al.* Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.* **6**, 705–714 (1992).
13. Borgel, J. *et al.* Targets and dynamics of promoter DNA methylation during early mouse development. *Nature Genet.* **42**, 1093–1100 (2010).
14. Meissner, A. Epigenetic modifications in pluripotent and differentiated cells. *Nature Biotechnol.* **28**, 1079–1088 (2010).
15. Razin, A. & Shemer, R. DNA methylation in early development. *Hum. Mol. Genet.* **4**, 1751–1755 (1995).
16. Monk, M., Boubelik, M. & Lahnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**, 371–382 (1987).
17. Rougier, N. *et al.* Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108–2113 (1998).
18. Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
19. Lane, N. *et al.* Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93 (2003).
20. Oswald, J. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
21. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182 (2002).
22. Kim, S. H. *et al.* Differential DNA methylation reprogramming of various repetitive sequences in mouse preimplantation embryos. *Biochem. Biophys. Res. Commun.* **324**, 58–63 (2004).
23. Bock, C. *et al.* Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nature Biotechnol.* **28**, 1106–1114 (2010).
24. Harris, R. A. *et al.* Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnol.* **28**, 1097–1105 (2010).
25. Smallwood, S. A. *et al.* Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genet.* **43**, 811–814 (2011).
26. Davis, T. & Vaisvila, R. High sensitivity 5-hydroxymethylcytosine detection in Balb/C brain tissue. *J. Vis. Exp.* **48** (2011).
27. Ficiz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–402 (2011).
28. Szulwach, K. E. *et al.* Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet.* **7**, e1002154 (2011).
29. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
30. Wu, H. *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684 (2011).
31. Xu, Y. *et al.* Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell* **42**, 451–464 (2011).
32. Branco, M. R., Ficiz, G. & Reik, W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature Rev. Genet.* **13**, 7–13 (2011).
33. Santos, F., Hendrich, B., Reik, W. & Dean, W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182 (2002).
34. Hajkova, P. *et al.* Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* **329**, 78–82 (2010).
35. Wossidlo, M. *et al.* Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes. *EMBO J.* **29**, 1877–1888 (2010).
36. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
37. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
38. Ichiyanagi, K. *et al.* Locus- and domain-dependent control of DNA methylation at mouse B1 retrotransposons during male germ cell development. *Genome Res.* **21**, 2058–2066 (2011).
39. Goodier, J. L., Ostertag, E. M., Du, K. & Kazanian, H. H. Jr. A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* **11**, 1677–1685 (2001).
40. Edwards, C. A. & Ferguson-Smith, A. Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**, 281–289 (2007).
41. Bestor, T. H. The DNA methyltransferases of mammals. *Hum. Mol. Genet.* **9**, 2395–2402 (2000).
42. Hirasawa, R. *et al.* Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* **22**, 1607–1616 (2008).
43. Lucifero, D. *et al.* Coordinate regulation of DNA methyltransferase expression during oogenesis. *BMC Dev. Biol.* **7**, 36 (2007).
44. Ramsahoye, B. H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA* **97**, 2373–2378 (2000).
45. Ziller, M. J. *et al.* Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* **7**, e1002389 (2011).
46. Haines, T., Rodenhiser, D. & Ainsworth, P. Allele-specific non-CpG methylation of the Nf1 gene during early mouse development. *Dev. Biol.* **240**, 585–598 (2001).
47. Tomizawa, S. *et al.* Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* **138**, 811–820 (2011).
48. Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E. & Spadafora, C. Expression of LINE-1 retrotransposons is essential for murine preimplantation development. *Mol. Reprod. Dev.* **73**, 279–287 (2006).
49. Kigami, D., Minami, N., Takayama, H. & Imai, H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol. Reprod.* **68**, 651–654 (2003).
50. Okada, Y., Yamagata, K., Hong, K., Wakayama, T. & Zhang, Y. A role for the elongator complex in zygotic paternal genome demethylation. *Nature* **463**, 554–558 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank all members of the Meissner and Reggev laboratories. M. Garber, N. Yosef, J. Ye, R. Koche, C. Bock, R. Maehr and D. Egli for technical advice and discussion. We thank all members of the Broad Sequencing Platform, in particular F. Kelly and J. Meldrim, T. Fennel, K. Tibbetts and J. Fostel. We also thank S. Levine, M. Gravina and K. Thai from the MIT BioMicro Center. A.R. is an investigator of the Merkin Foundation for Stem Cell Research at the Broad Institute. This work was supported by the NIH Pioneer Award (5DP1OD003958), the Burroughs Wellcome Career Award at the Scientific Interface and HHMI (to A.R.), the Harvard Stem Cell Institute (to T.S.M.) and the NIH (5RC1AA019317, U01ES017155 and P01GM099117), the Massachusetts Life Science Center and the Pew Charitable Trusts (to A.M.) and a Center for Excellence in Genome Science from the NHGRI (1P50HG006193-01, to A.R. and A.M.).

Author Contributions Z.D.S. and A.M. conceived the study and Z.D.S., M.M.C. and A.M. facilitated its design. Z.D.S. collected samples and performed methylation profiling. M.M.C. performed all analysis with assistance from T.S.M. and Z.D.S. H.G. and A.G. provided critical technical assistance and expertise. Z.D.S., M.M.C., T.S.M., A.R. and A.M. interpreted the data. Z.D.S., M.M.C. and A.M. wrote the paper with the assistance of the other authors.

Author Information RRBS data is deposited at the Gene Expression Omnibus under accession number GSE34864. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.M. (alexander_meissner@harvard.edu).

METHODS

Preparation of samples. Isolation of gametes, pre- and post-implantation embryos was performed using procedures described in detail elsewhere⁵¹. Briefly, 4–6-week-old BDF1 female mice (Charles River) were injected with 5 IU of pregnant mare gonadotropin (Sigma) followed 46 h later by 5 IU human chorionic gonadotropin (Sigma). Primed mice were then either directly used to collect oocytes or mated with 129X1 male mice (Jackson) to collect fertilized embryos. Twelve hours after final hormone injection, oocytes or zygotes were isolated from the ampulla under mineral oil and collected in hyaluronidase-containing M2 medium (Millipore) drops to eliminate cumulus cells or sperm contaminants. Oocytes were then depleted of somatic contaminants via progressive dilution through sequential drops of CO₂ buffered, amino-acid-supplemented KSOM medium (Millipore) until no somatic contaminants were observed.

Embryos were cultured in KSOM until collection at progressive cleavage stages with isolation occurring within 6 h of the first observed cleavage event for that stage. Zygotes were screened for the presence of visible pronuclei and subjected to XY Clone (Hamilton Thorne) laser-assisted polar body biopsy using an 8- μ m bore piezo pipette (Humagen, Supplementary Fig. 1 and Supplementary Movie 1). Clean cleavage stage embryos underwent an identical approach, with developmental progression unhindered by biopsy conducted at the 2-cell stage (Supplementary Fig. 1). For each collection, batches of embryos were carefully screened to ensure that each stage did not contain any abnormal embryos. Collection for zygotes was timed at \sim 10 h.p.f. with fertilization assumed to occur 6–8 h after HCG injection, which was again confirmed by the relative synchronicity of the first cleavage division and by relative pronuclear stage. Biopsies were conducted in M2 media (Millipore) in batches of 5–10 embryos to reduce time on the micro-manipulator stage. Before the final collection, cleaned and sorted samples were washed with acid Tyrode's solution (Sigma) to eliminate the zona pellucida and to deplete any residual somatic contaminants or polar bodies through a short series of additional washes.

Cavitated E3.5 blastocysts were flushed from the uteri of naturally mated mice into M2 or DMEM followed by sequential washing in KSOM. The ICM itself was enriched from collected blastocysts by treating the embryo with rabbit anti-mouse serum (Sigma) before immunosurgical depletion of the trophectoderm using guinea-pig complement serum (Sigma). Isolated ICMs were serially washed to remove contaminants.

E6.5 and 7.5 embryos were isolated after mechanical dissection of the decidua from the uterine lining of mated mice. Samples were again serially washed and peripheral trophectodermal tissues dissected away using fine glass capillaries.

Swimming sperm samples were isolated in BSA-supplemented human tubule fluid (Millipore) from the caudal epididymis of male mice within 5 days of a successful natural mating as scored by copulation plug. Sperm were incubated in buffered HTF as in *in vitro* fertilization for over 1 h in part to reduce somatic contaminants, and samples were scored for relative quality under \times 10 microscopy before snap freezing.

All samples were then collected at minimal volume and either snap frozen or immediately re-suspended in DNA lysis buffer.

Preparation of reduced representation bisulphite sequencing libraries. RRBS libraries were generated as described previously^{52–54}. Briefly, DNA was isolated from snap-frozen embryos in DNA lysis buffer (100 mM Tris-HCl (pH 8.5), 5 mM EDTA, 0.2% SDS, 200 mM NaCl) supplemented with 300 μ g ml⁻¹ proteinase K (Invitrogen) followed by phenol:chloroform extraction, ethanol precipitation and re-suspension in EB buffer. Isolated DNA was then subjected to MspI digestion (NEB), end repair using Klenow 3'-5' exo- (NEB) supplemented with GTP, meCTP and ATP in a 1:1:10 ratio to facilitate 3' A tailing, and ligation of standard adapters using ultraconcentrated 10⁶ U T4 DNA ligase (NEB) and extended 20 h ligation at 16 °C. Size selection of 40–150 and 150–270-bp fragments containing ligated adaptor was conducted by extended gel electrophoresis using NuSieve 3:1 agarose (Lonza) and gel extraction (Qiagen) using 50 ng dephosphorylated, sonicated *Escherichia coli* DNA as a protective carrier and to increase overall yield. The isolated molecular weight fractions in a given RRBS library were then separately treated with sodium bisulphite using the EpiTect bisulphite conversion and column purification system (Qiagen) with a modified conversion strategy as described⁵². After clean up, the optimal, minimum PCR cycle number required to generate the final libraries was gauged using diagnostic PCRs for each library. Final libraries were then generated from the complete bisulphite converted pool and purified through a second round of gel electrophoresis. High- and low-molecular-mass fragments were then either sequenced separately or pooled at a 2:1 ratio by mass to assume an equimolar representation of both size ranges. Libraries were then sequenced on an Illumina Genome Analyzer II before alignment and analysis. The sequencing reads were aligned to the Mouse Genome Build 37 (mm9) using a custom computational pipeline taking into account the

strain background for each sample^{4,54}. To supplement our data set we included sperm replicate 2 from ref. 25 (SRA accession ERP000689).

Estimating methylation levels. The methylation level of each sampled cytosine was estimated as the number of reads reporting a C, divided by the total number of reads reporting a C or T. Single CpG methylation levels were limited to those CpGs that had at least tenfold coverage. For 100-bp tiles, reads for all the CpGs that were covered more than fivefold within the tile were pooled and used to estimate the methylation level as described for single CpGs. The CpG density for a given single CpG is the number of CpGs 50 bp up- and downstream of that CpG. The CpG density for a 100-bp tile is the average of the CpG density for all single CpGs used to estimate methylation level in the tile. CpA methylation levels were estimated in the same way as for CpG methylation.

The methylation level reported for a sample is the average methylation level across replicates. A replicate will contribute to the average only if it meets the coverage criteria within the replicate. Technical replicates were averaged before contributing to the sample average.

Genomic features. High-density CpG promoters (HCP), intermediate-density CpG promoters (ICP), low-density CpG promoters (LCP), TSS, CpG island, and DMR annotations were taken from ref. 4. Promoters are defined as 1 kb up- and downstream of the TSS. LINE, LTR and SINE annotations were downloaded from the UCSC browser (mm9) RepeatMasker tracks. Gene annotations were downloaded from the UCSC browser (mm9) RefSeq track. In each case, the methylation level of an individual feature is estimated by pooling read counts for all CpGs within the feature that are covered greater than fivefold, and levels are only reported if a feature contains at least 5 CpGs with such coverage (in contrast to 100-bp tiles where no minimum number of CpGs is required). A tile is annotated as a genomic feature if any portion of the tile overlaps with the feature and may be annotated by more than one feature (for example, the same tile can be annotated as both a promoter and a gene).

Identification of tiles with changing methylation levels and their enrichments.

A tile is considered changing if it both has a methylation difference \geq 0.2 between two stages and is significant in a two-sample *t*-test with unequal variance after correction for multiple hypothesis testing (FDR < 0.05) using the Benjamini-Hochberg method⁵⁵. Enrichment *P* values are from the hypergeometric distribution where the background is the number of tiles that have a methylation difference \geq 0.2 and are corrected for multiple hypotheses at FDR < 0.05, based on the number of feature sets tested.

Identification of enriched retrotransposon families. The same procedure for identifying changing tiles was applied to the methylation levels of retrotransposon elements to identify changing elements. Enrichment for families was done using annotations from the RepeatMasker track of the UCSC genome browser.

Novel DMR identification. 100-bp tiles where one gamete had a mean methylation greater than 0.75 and the other gamete had a mean methylation of less than 0.25 were flagged as potential DMRs. Linear regression was used to identify tiles which had methylation levels in zygote which were halfway between the methylation levels in oocyte and sperm. Only tiles that had two replicates present in each time point were considered. Residuals were calculated as the mean of the differences between the model predictions and the data taking into account missing values. ANOVA was used to assign a *P* value to each tile. A tile was considered a novel DMR if it had a residual in the tenth percentile of tiles tested and a significant *P* value from ANOVA with a Benjamini-Hochberg FDR < 0.05. A residual in the tenth percentile corresponds to an FDR < 0.1 by a permutation test where zygote methylation values are shuffled for potential DMR tiles. In the pie charts (Fig. 5b, d), the genomic feature that covered the most novel tiles was reported first and then subtracted from the set before reporting the feature which covered the next largest number of tiles. This procedure was repeated until all tiles were categorized. The one exception was for oocyte-contributed DMRs where promoters were taken out before genes.

Identification of SNPs. An initial set of SNPs between 129X1 and BDF1 (C57BL/6 \times DBA/2) was taken from Mouse Genome Informatics⁵⁶. The set was filtered such that SNPs that fell into the following categories were removed: (1) SNPs that had inconsistent entries for the same position; (2) SNPs not trackable by RRBS (C/T or A/G); (3) SNPs between C57BL/6 and DBA/2; and (4) SNPs that were not covered by X1 and BDF1 in an *in silico* digest. The log odds ratio [$\log_2(X1 \text{ count} + 0.01 / C57 \text{ count} + 0.01)$] was calculated for each SNP that was covered in the data set ($n = 786$). SNPs that had at least 10 \times coverage with an average log odds ratio across all replicates between -5 and 5 and a sperm X1 log odds ratio greater than 2 were considered to be of stringent quality ($n = 636$) and used to assess both maternal bias and to serve as a general quality control metric for all libraries incorporated into the data set.

Parent-of-origin methylation tracking. The 636 SNPs identified above corresponded to 1,674 CpG dinucleotides and were used to track allelic single CpG methylation. Reads were segregated into either X1 or BDF1 backgrounds according

to SNP type, and CpG methylation levels were called in the same manner described above. SNP normalized methylation values (Supplementary Fig. 4) are the average of the methylation values from each strain.

51. Nagy, A. *Manipulating the Mouse Embryo: a Laboratory Manual* 3rd edn (Cold Spring Harbor Laboratory Press, 2003).
52. Smith, Z. D., Gu, H., Bock, C., Gnirke, A. & Meissner, A. High-throughput bisulfite sequencing in mammalian genomes. *Methods* **48**, 226–232 (2009).
53. Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols* **6**, 468–481 (2011).
54. Gu, H. *et al.* Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nature Methods* **7**, 133–136 (2010).
55. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
56. Blake, J. A., Bult, C. J., Kadin, J. A., Richardson, J. E. & Eppig, J. T. The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. *Nucleic Acids Res.* **39**, D842–D848 (2011).

Crystal structure of a membrane-embedded H^+ -translocating pyrophosphatase

Shih-Ming Lin¹, Jia-Yin Tsai², Chwan-Deng Hsiao², Yun-Tzu Huang¹, Chen-Liang Chiu¹, Mu-Hsuan Liu¹, Jung-Yu Tung¹, Tseng-Huang Liu¹, Rong-Long Pan¹ & Yuh-Ju Sun¹

H^+ -translocating pyrophosphatases (H^+ -PPases) are active proton transporters that establish a proton gradient across the endomembrane by means of pyrophosphate (PP_i) hydrolysis^{1,2}. H^+ -PPases are found primarily as homodimers in the vacuolar membrane of plants and the plasma membrane of several protozoa and prokaryotes^{2,3}. The three-dimensional structure and detailed mechanisms underlying the enzymatic and proton translocation reactions of H^+ -PPases are unclear. Here we report the crystal structure of a *Vigna radiata* H^+ -PPase (VrH^+ -PPase) in complex with a non-hydrolysable substrate analogue, imidodiphosphate (IDP), at 2.35 Å resolution. Each VrH^+ -PPase subunit consists of an integral membrane domain formed by 16 transmembrane helices. IDP is bound in the cytosolic region of each subunit and trapped by numerous charged residues and five Mg^{2+} ions. A previously undescribed proton translocation pathway is formed by six core transmembrane helices. Proton pumping can be initialized by PP_i hydrolysis, and H^+ is then transported into the vacuolar lumen through a pathway consisting of Arg 242, Asp 294, Lys 742 and Glu 301. We propose a working model of the mechanism for the coupling between proton pumping and PP_i hydrolysis by H^+ -PPases.

Two proton-pumping proteins, vacuolar H^+ -ATPases (V-ATPases) and H^+ -PPases, coexist on plant vacuolar membranes and use ATP and PP_i , respectively, as energy sources for H^+ translocation⁴. Both of these enzymes acidify the vacuolar lumen and establish an electrochemical proton gradient across the vacuolar membrane. Whereas V-ATPases have been widely studied, the structure and function of H^+ -PPases are as yet unknown. The H^+ -PPase was first identified from photosynthetic bacterium *Rhodospirillum rubrum*⁵, after which the functional and biochemical investigations of vacuolar H^+ -PPases from plants were reported^{6,7}. Subsequently, the protein chemical identification of H^+ -PPases from red beet⁸ and mung bean⁹ and the molecular cloning of the enzyme from *Arabidopsis*¹⁰ were gradually completed. Afterwards, the heterologous expression of H^+ -PPase in yeast was established and the sufficiency of PP_i -energized H^+ translocation was demonstrated¹¹. On the basis of these arduous results, a fourth category of primary H^+ pump, H^+ -PPases, was thus determined¹².

H^+ -PPases have a high degree of amino acid sequence homology (86–91% identity in land plants; Supplementary Fig. 1)², and they can be divided into two subfamilies: type I (K^+ -dependent) and type II (K^+ -independent)¹³. In *Arabidopsis*, the H^+ -PPase regulates organ development by modulating the apoplastic pH and eliminating the metabolic by-product, cytosolic PP_i (refs 14, 15). The overexpression of H^+ -PPases improves tolerance to salinity and drought in many higher plants by increasing ion retention^{16,17}. A PP_i -binding motif, (E/D)(X)₇KXE, and two acidic motifs, DX₃DX₃D, have been suggested as being essential for the enzymatic function of H^+ -PPases^{2,12,18}. Nevertheless, the three-dimensional structure of H^+ -PPase is unknown. We therefore sought to solve its atomic structure to gain a better understanding of the machinery of this unique proton pump.

We determined the crystal structure of VrH^+ -PPase bound to IDP by the multiple-wavelength anomalous dispersion (MAD) method and multiple isomorphous replacement with anomalous scattering (MIRAS) (Supplementary Table 1). The protein is compactly folded in a rosette manner with 16 transmembrane helices (TMs M1–M16) within two concentric walls (Fig. 1a). Six TMs (M5, M6, M11, M12, M15 and M16) bundle together to form an inner wall, which is in turn surrounded by ten additional TMs (M1–M4, M7–M10 and M13–M14) that constitute the outer wall (Fig. 1b). The TMs fold sequentially counterclockwise in the inner and outer walls with the exception of M1, which tilts slightly outward. There are two short helices ($\alpha 2$ and $\alpha 3$) on the cytosolic side, and two additional helices ($\alpha 1$ and $\alpha 4$) together with two antiparallel β strands ($\beta 1$ and $\beta 2$) on the luminal side (Supplementary Fig. 2). A disulphide bond (C124–C132) was detected in loop 2 (Supplementary Fig. 2). Both the amino and carboxy termini face the vacuolar lumen (Fig. 1c). Each subunit contains a single IDP molecule embedded in the core of the inner wall near the cytosolic region. It is conceivable that the inner wall is responsible for substrate binding, and the outer wall probably maintains the structural integrity of the protein.

VrH^+ -PPase is a homodimer with an extensively buried 3,241 Å² surface between the subunits (Fig. 1c). The electrostatic surface potential of the VrH^+ -PPase dimer is shown in Fig. 1d. The cytosolic region of VrH^+ -PPase has many hydrophilic residues, whereas the vacuolar region of the protein protruding out of the membrane is smaller. The homodimer has two-fold symmetry and a root mean squared deviation of 0.32 Å among the C α atoms between the two subunits (Supplementary Fig. 3a). Four TMs (M5, M12, M15 and M16) from the inner wall and two TMs (M10 and M13) from the outer wall participate in subunit interactions that primarily include hydrophobic interactions, six hydrogen bonds and two salt bridges (Supplementary Fig. 3b–d and Supplementary Table 2).

The substrate/IDP-binding site is a funnel-shaped pocket formed by six core TMs with a solvent-accessible volume of 1,521 Å³ (Fig. 2a). The electrostatic surface potential of the IDP-binding pocket is shown in Fig. 2b. The pocket has an unusually acidic environment that contains 12 acidic residues (Asp 253, Asp 257, Glu 268, Asp 269, Asp 279, Asp 283, Asp 287, Asp 507, Asp 691, Asp 723, Asp 727 and Asp 731). Three lysine residues (Lys 250, Lys 730 and Lys 694) and one asparagine residue (Asn 534) are also found in this pocket (Fig. 2c). All these residues are highly conserved among H^+ -PPases (Supplementary Fig. 1). Site-directed mutagenesis studies indicate that most of these conserved residues are essential for PP_i hydrolysis^{18–20}. In the current structure, the IDP molecule was bound directly by three lysine residues through hydrogen bonds. In addition, one K^+ ion and five Mg^{2+} ions (Mg1–Mg5) were identified around the IDP molecule at the pocket (Fig. 2c). K^+ is the stimulator of the type I H^+ -PPases¹, and Mg^{2+} is essential for the activity of H^+ -PPases²¹. All these Mg^{2+} ions mediate the interactions between IDP and Asp/Asn residues. The binding interactions (Supplementary Table 3) precisely confine the

¹Department of Life Science and Institute of Bioinformatics and Structural Biology, College of Life Science, National Tsing Hua University, Hsinchu 30013, Taiwan. ²Institute of Molecular Biology, Academia Sinica, Taipei 11529, Taiwan.

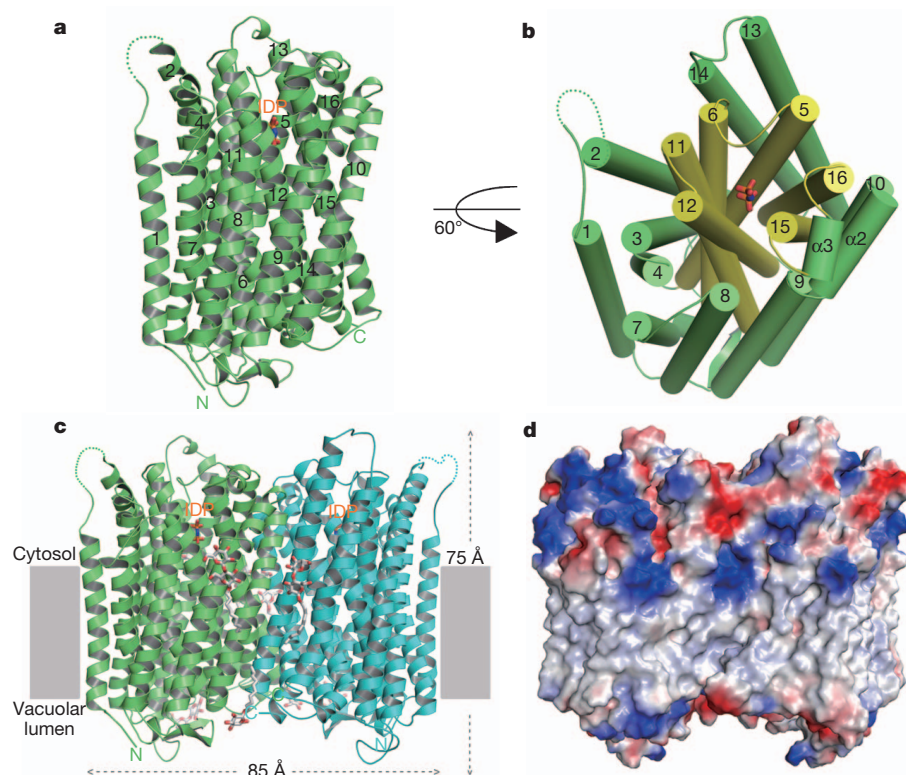


Figure 1 | The crystal structure of the VrH^+ -PPase-IDP complex. **a**, Ribbon diagram of the overall structure of VrH^+ -PPase, containing 16 TMs (labelled 1–16). A missing region (residues 42–66) is shown with dotted lines. **b**, The six inner and ten outer TMs drawn as cylinders and coloured in yellow and green, respectively. This orientation is rotated by 60° from that in **a**. **c**, VrH^+ -PPase

dimer shown as a ribbon diagram with height and width dimensions of 75 Å and 85 Å, respectively. The detergent molecules of n-decyl β -D-maltoside are shown as sticks. **d**, Electrostatic surface potential of the VrH^+ -PPase dimer (red, blue and white indicate negative, positive and neutral potentials, respectively). In **a**–**c**, IDP is shown as sticks and coloured in CPK.

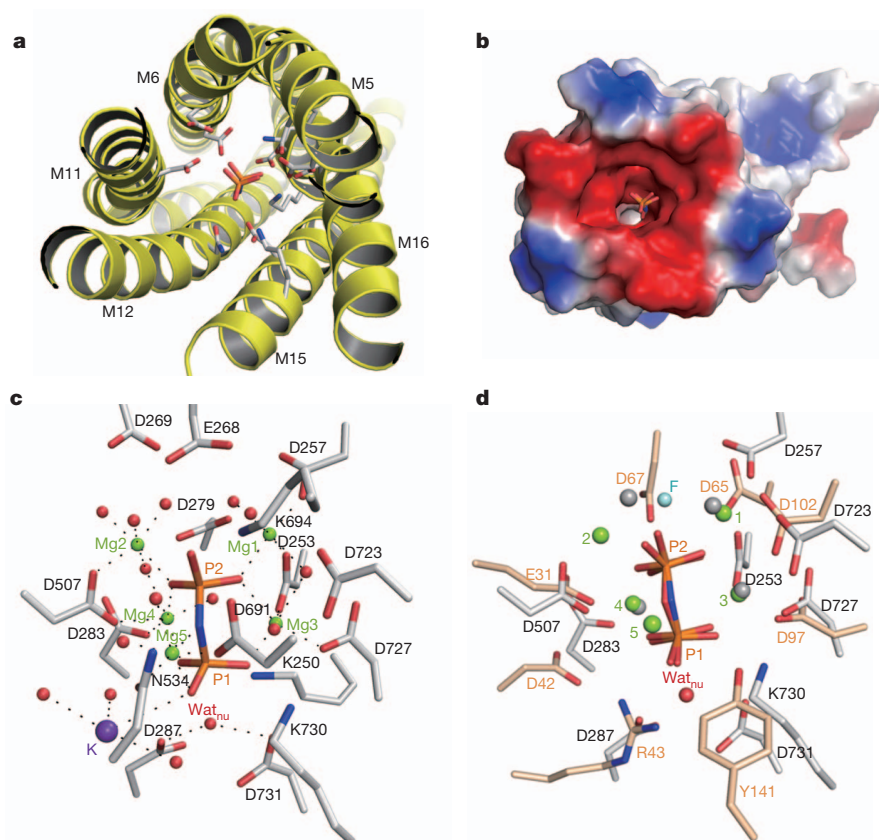


Figure 2 | The substrate-binding site. **a**, Six core TMs (yellow ribbon) with IDP-binding residues (sticks). **b**, The electrostatic surface potential of the IDP binding pocket (red, blue and white indicate negative, positive and neutral potentials, respectively). **c**, The IDP-binding residues. Mg^{2+} ions, K^+ ions and water molecules are shown as green, purple and red spheres, respectively. Interactions are presented as dashed lines. **d**, The binding site of VrH^+ -PPase (in white) superimposed on *EcPPase* (PDB 2AUU; in pink). The Mg^{2+} ions of the VrH^+ -PPase (in green, with numbers) and *EcPPase* (in grey) are shown as spheres. The F^- in *EcPPase* is shown as a blue sphere, and the Wat_{nu} in VrH^+ -PPase is shown as a labelled red sphere. IDP in VrH^+ -PPase and PP_i in *EcPPase* are coloured in CPK.

substrate at the active site in a proper orientation for enzymatic hydrolysis.

The current structure indicates that the highly conserved fragment D253-X₃-D257-X₃-K261-X-E263 from M5, corresponding to the DX₇KXE motif, is involved in substrate binding^{12,18}. Both Asp 253 and Asp 257 interact with IDP through Mg1 and Mg3 (Fig. 2c). However, Lys 261 and Glu 263 do not participate directly in IDP binding. Lys 261 contributes to a salt-bridge network (Supplementary Fig. 4 and Supplementary Table 4) that connects the core TMs to stabilize the active site. Glu 263 forms a salt bridge with Arg 609 and probably contributes to the structural stability of V_rH⁺-PPase. Furthermore, the highly conserved acidic motifs D279-X₃-D283-X₃-D287 and D723-X₃-D727-X₃-D731 as well as all aspartic acids, except Asp 287 and Asp 731, interact with IDP through either Mg²⁺ ions or water.

In Fig. 2d the substrate and the surrounding residues in V_rH⁺-PPase are superimposed on those of type I soluble *Escherichia coli* pyrophosphatase (EcPPase)²². Both enzymes contain several aspartic residues for coordinating Mg²⁺ ions to bind the substrate at the active site. There are five and four Mg²⁺ ions in the binding pockets of V_rH⁺-PPase and EcPPase, respectively. The binding of Mg5 to IDP is detected only in V_rH⁺-PPase. The locations of Mg3 and Mg4 in V_rH⁺-PPase are similar to those of the corresponding Mg²⁺ ions in EcPPase. In contrast, Mg1 and Mg2 are located at distinct positions in both enzymes. The distances between Mg1 and Mg2 in V_rH⁺-PPase and EcPPase are 5.3 and 3.5 Å, respectively. In EcPPase, Mg1 and Mg2 bind with a specific inhibitor, F⁻, which occupies the position of the nucleophile close to the P2 phosphate of PP_i, to prevent hydrolysis of the substrate²². However, a similar binding phenomenon is not observed in V_rH⁺-PPase, because the distance between Mg1 and Mg2 is too large to bind a nucleophile for attacking PP_i (Fig. 2d). Instead, in V_rH⁺-PPase, a water molecule (Wat_{nu}) is found near the P1 phosphate of IDP (2.6 Å) and forms hydrogen bonds with Asp 287 and Asp 731, respectively. It is conceivable that Wat_{nu} might act as a nucleophile for PP_i hydrolysis in V_rH⁺-PPase. Asp 287 and Asp 731 are highly conserved among H⁺-PPases but are replaced by Arg 43 and Tyr 141 in soluble EcPPase. F⁻ cannot compete with Wat_{nu} for the same binding site, presumably as a result of the repulsion between negatively charged F⁻ and aspartic acid. H⁺-PPase is therefore less sensitive to F⁻ inhibition than soluble PPase is². These differences suggest that membrane-bound H⁺-PPases and soluble PPases use different strategies to trap the nucleophile for PP_i hydrolysis.

Previous evidence suggests that the dimerization of H⁺-PPase is important for H⁺ translocation activity²³. However, the possible residues that can be protonated as part of proton conduction could not be found in the dimer interface in V_rH⁺-PPase. Instead, four lined-up charged residues, Arg 242, Asp 294, Lys 742 and Glu 301, were observed in a narrow, compact and water-inaccessible transmembrane region of the core TMs (Fig. 3a). These residues are highly conserved among H⁺-PPases except Glu 301, which is found only in higher plants (Supplementary Fig. 1). Glu 301 is located at a rather narrow point in the pathway and close to the vacuolar lumen. It might therefore act as a constricting neck (Fig. 3a). Mutations at Glu 301 abolish proton translocation activity (Supplementary Table 5)^{19,24,25}. In addition, Glu 301 contributes to the binding of the lipophilic carbodiimide *N,N'*-dicyclohexylcarbodiimide (DCCD), a potent blocker of proton conductance²⁴. These data indicate that Glu 301 probably serves as the proton donor/acceptor in the proton translocation mechanism. Furthermore, Asp 294 and Lys 742 form the only salt bridge among these acid–base pairs (Fig. 3a). Lys 742 might reciprocally modulate the protonation and deprotonation of Asp 294 and Glu 301, similar to that in the regulatory machinery of bacteriorhodopsin²⁶. Arg 242, which is located near the PP_i-binding site, may form a salt bridge with Asp 287 or Asp 731, resulting in the deprotonation of the latter residues. Site-directed mutagenesis studies revealed that Arg 242, Asp 294 and Lys 742 are crucial for PP_i hydrolysis as well as H⁺ translocation (Supplementary Table 5)^{19,27}. Together, these acid–base pairs,

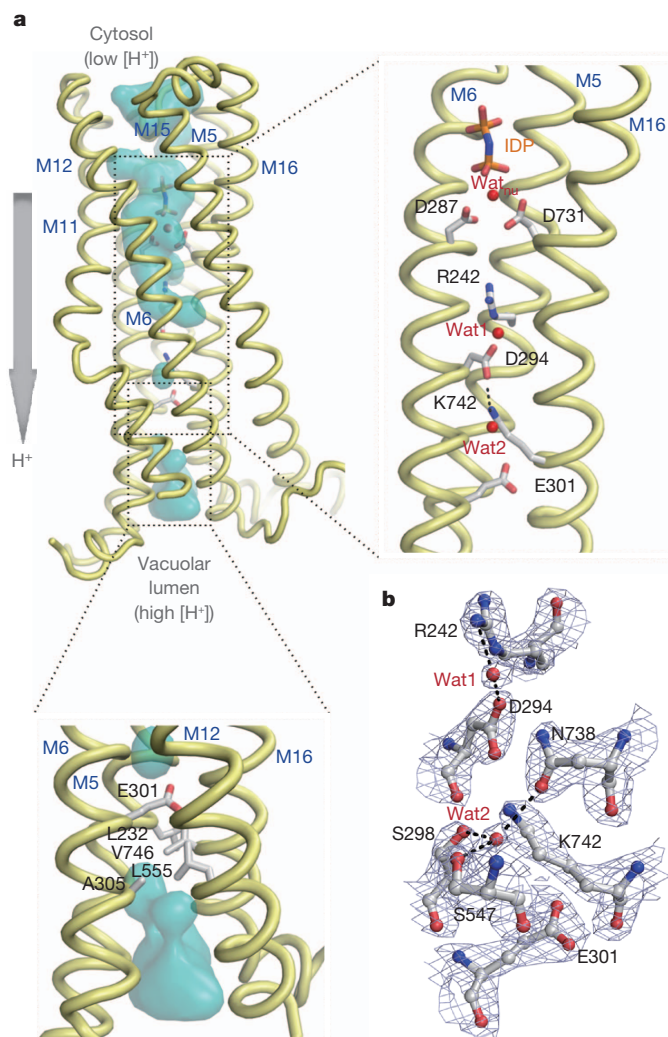


Figure 3 | The proton transport pathway of V_rH⁺-PPase. **a**, The proton transport pathway is formed by six core TMs (M5, M6, M11, M12, M15 and M16). The solvent-accessible surface area is coloured in cyan. The arrow indicates the direction of proton translocation. Right: zoomed-in view of the proton transport pathway. The residues involved in proton transport are shown and labelled. The solvent-accessible surface has been removed. Bottom: the hydrophobic gate around Glu 301. Residues forming a hydrophobic gate are displayed and labelled. **b**, The electron density map ($2F_{\text{obs}} - F_{\text{calc}}$) (in blue) around the proton transport pathway drawn at a contour level of 2σ . The IDP is shown as sticks and coloured in CPK. Water molecules Wat_{nu}, Wat1 and Wat2 are presented as labelled red spheres. Water-mediated hydrogen bonds are drawn as black dashed lines.

Arg 242, Asp 294, Lys 742 and Glu 301, form a potential proton transport pathway surrounded by six core TMs that convey protons from the active site to the vacuolar lumen.

The proton-pumping proteins usually use bound water molecules to assist proton transport²⁸; examples are bacteriorhodopsin²⁶ and P-type H⁺-ATPases²⁹. Nevertheless, two ordered water molecules (Wat1 and Wat2) along the proton transport pathway of V_rH⁺-PPase were identified from the electron density map (Fig. 3b). Wat1 forms hydrogen bonds with Arg 242 and Asp 294 and seems to be a continuation of the bulk solvent at the PP_i-binding site. Wat2 was found in the vicinity of Lys 742 (with a distance of 3.7 Å) and trapped by Asp 294, Ser 298, Ser 547, Asn 738 and Lys 742 with hydrogen bonding to Ser 298, Ser 547 and Asn 738. The bound waters, Wat1 and Wat2, presumably facilitate proton transport in the pathway of V_rH⁺-PPase.

In the V_rH⁺-PPase-IDP complex, the vacuolar portion of the proton transport pathway is relatively narrow and is occupied by hydrophobic residues, such as Leu 232 (M5), Ala 305 (M6), Leu 555 (M12) and

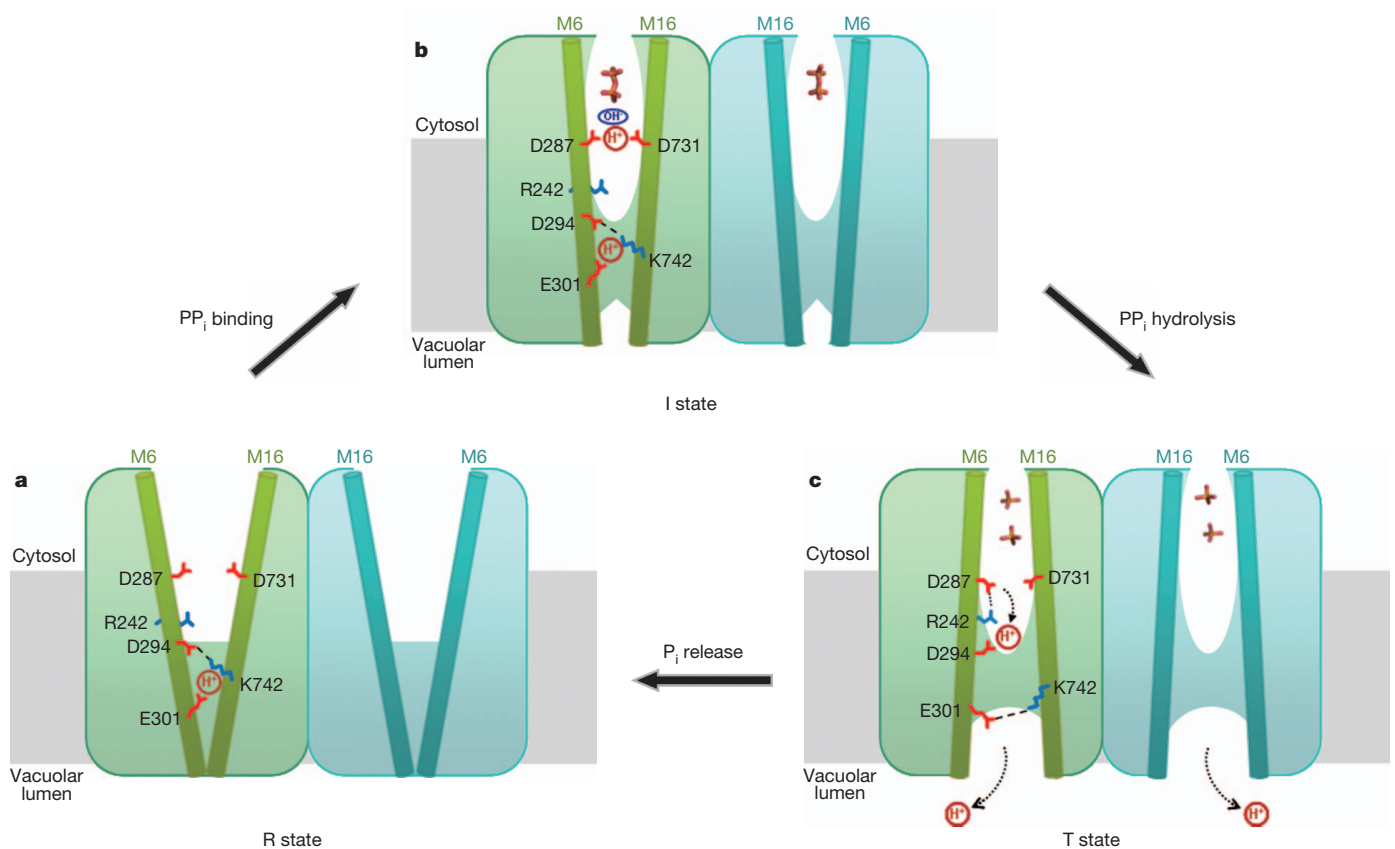


Figure 4 | A working model for proton pumping in VrH^+ -PPase. **a**, Resting state (R state). **b**, Initiated state (I state). **c**, Transient state (T state). The VrH^+ -PPase dimer is shown and coloured in green and blue. M6 and M16 are shown

Val 746 (M16) (Fig. 3a). Most of these residues are conserved among H^+ -PPases from various species (Supplementary Fig. 1). They possibly form a hydrophobic gate, keeping Glu 301 away from the hydrophilic environment of the vacuolar lumen, where the H^+ concentration is higher. Such a gate prevents the H^+ from refluxing and maintains unidirectional H^+ translocation from cytosol to lumen. It is conceivable that the constricted pathway and the acid–base pairs have crucial functions in directional proton pumping of the H^+ -PPase.

On the basis of previous mutational and biochemical studies (Supplementary Table 5) and the present structural findings, we propose a working model for the coupling of PP_i hydrolysis to proton pumping (Fig. 4). In this model, the process involves three sequential states: R, I and T. Similarly, the limited trypsinolysis analysis of H^+ -PPase^{20,21} (Supplementary Fig. 5) indicates that the binding pocket of H^+ -PPase is flexible and might be exposed to the solvent in the absence of the substrate, and the protein is in a resting state (R state; Fig. 4a). In addition, an open conformation at the active site in the absence of substrate (PP_i) was also supported by a study using single-molecule fluorescence resonance energy transfer (FRET)³⁰. In contrast, the luminal portion of the core TMs assumes a closed conformation that prevents H^+ from back-flushing into the cytosol. On binding of substrate (or IDP in our structure), H^+ -PPase enters an initiated state for PP_i hydrolysis (I state; Fig. 4b). The core TMs then change into a closed conformation on the cytosolic side, locking PP_i in the substrate-binding pocket. At this stage, the luminal portion of these TMs is changed into a semi-open conformation for sequential H^+ translocation. The structure of the VrH^+ -PPase–IDP complex that we have resolved would represent the protein in the I state. On hydrolysis of PP_i , free phosphates are generated with concomitant proton production (Fig. 4c). Then the proton relay through Arg 242–Asp 294–Lys 742–Glu 301 occurs as a chain reaction, and the core TMs adopt

as cylinders. The residues involved in proton transport are shown and labelled. PP_i and free P_i are shown as sticks and coloured in CPK. The proton and hydroxyl ions are labelled. The possible salt bridges are shown as dashed lines.

an open conformation on the vacuolar lumen side for proton release. Consequently, H^+ -PPase proceeds to the transient state (T state; Fig. 4c). After the release of P_i to the cytosol and translocation of a proton to the vacuolar lumen, H^+ -PPase returns to the R state. This proton-pumping mechanism is completed through a series of delicate conformational changes driven primarily by the energy derived from PP_i hydrolysis. Thus, our working model provides a scheme capable of accounting for how PP_i hydrolysis and the movement of protons from one side of the membrane to the other might be accomplished by the H^+ -PPase.

METHODS SUMMARY

Vigna radiata H^+ -translocating pyrophosphatase (VrH^+ -PPase) fused with a C-terminal His-tag was expressed and isolated from *Saccharomyces cerevisiae* with methods that have been described previously. In brief, VrH^+ -PPase was crystallized by using 23% (w/v) PEG2000 precipitant in 50 mM MES pH 6.5. The structural phase was determined from the tantalum derivative Ta_6Br_{12} by MAD and also from OsO_4 and Pt-organic derivatives by MIRAS (Supplementary Table 1). Model building was performed with COOT, and the final model was refined with REFMAC5 to an R -factor of 16.80% and an R_{free} of 20.31% at 2.35 Å. The refinement statistics are summarized in Supplementary Table 1.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 September 2011; accepted 17 February 2012.

Published online 28 March 2012.

1. Rea, P. A. & Poole, R. J. Vacuolar H^+ -translocating pyrophosphatase. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**, 157–180 (1993).
2. Maeshima, M. Vacuolar H^+ -pyrophosphatase. *Biochim. Biophys. Acta* **1465**, 37–51 (2000).
3. Serrano, A., Perez-Castineira, J. R., Baltscheffsky, H. & Baltscheffsky, M. Proton-pumping inorganic pyrophosphatases in some archaea and other extremophilic prokaryotes. *J. Bioenerg. Biomembr.* **36**, 127–133 (2004).

4. Maeshima, M. Tonoplast transporters: organization and function. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 469–497 (2001).
5. Baltscheffsky, H., Von Stedingk, L. V., Heldt, H. W. & Klingenberg, M. Inorganic pyrophosphate: formation in bacterial photophosphorylation. *Science* **153**, 1120–1122 (1966).
6. Chanson, A., Fichmann, J., Spear, D. & Taiz, L. Pyrophosphate-driven proton transport by microsomal membranes of corn coleoptiles. *Plant Physiol.* **79**, 159–164 (1985).
7. Rea, P. A. & Poole, R. J. Proton-translocating inorganic pyrophosphatase in red beet (*Beta vulgaris* L.) tonoplast vesicles. *Plant Physiol.* **77**, 46–52 (1985).
8. Britten, C. J., Turner, J. C. & Rea, P. A. Identification and purification of substrate binding subunit of higher plant H^+ -translocating inorganic pyrophosphatase. *FEBS Lett.* **256**, 200–206 (1989).
9. Maeshima, M. & Yoshida, S. Purification and properties of vacuolar membrane proton-translocating inorganic pyrophosphatase from mung bean. *J. Biol. Chem.* **264**, 20068–20073 (1989).
10. Sarafian, V., Kim, Y., Poole, R. J. & Rea, P. A. Molecular cloning and sequence of cDNA encoding the pyrophosphate-energized vacuolar membrane proton pump of *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **89**, 1775–1779 (1992).
11. Kim, E. J., Zhen, R. G. & Rea, P. A. Heterologous expression of plant vacuolar pyrophosphatase in yeast demonstrates sufficiency of the substrate-binding subunit for proton transport. *Proc. Natl Acad. Sci. USA* **91**, 6128–6132 (1994).
12. Rea, P. A. *et al.* Vacuolar H^+ -translocating pyrophosphatases: a new category of ion translocase. *Trends Biochem. Sci.* **17**, 348–353 (1992).
13. Belogurov, G. A. & Lahti, R. A lysine substitute for K^+ . A460K mutation eliminates K^+ dependence in H^+ -pyrophosphatase of *Carboxydotherrhus hydrogeniformans*. *J. Biol. Chem.* **277**, 49651–49654 (2002).
14. Li, J. *et al.* *Arabidopsis* H^+ -PPase AVP1 regulates auxin-mediated organ development. *Science* **310**, 121–125 (2005).
15. Ferjani, A. *et al.* Keep an eye on PP: the vacuolar-type H^+ -pyrophosphatase regulates postgerminative development in *Arabidopsis*. *Plant Cell* **23**, 2895–2908 (2011).
16. Park, S. *et al.* Up-regulation of a H^+ -pyrophosphatase (H^+ -PPase) as a strategy to engineer drought-resistant crop plants. *Proc. Natl Acad. Sci. USA* **102**, 18830–18835 (2005).
17. Guo, S. *et al.* Molecular cloning and characterization of a vacuolar H^+ -pyrophosphatase gene, SsVP, from the halophyte *Suaeda salsa* and its overexpression increases salt and drought tolerance of *Arabidopsis*. *Plant Mol. Biol.* **60**, 41–50 (2006).
18. Nakanishi, Y., Saijo, T., Wada, Y. & Maeshima, M. Mutagenic analysis of functional residues in putative substrate-binding site and acidic domains of vacuolar H^+ -pyrophosphatase. *J. Biol. Chem.* **276**, 7654–7660 (2001).
19. Hirono, M., Nakanishi, Y. & Maeshima, M. Essential amino acid residues in the central transmembrane domains and loops for energy coupling of *Streptomyces coelicolor* A3(2) H^+ -pyrophosphatase. *Biochim. Biophys. Acta* **1767**, 930–939 (2007).
20. Lee, C. H. *et al.* Identification of essential lysines involved in substrate binding of vacuolar H^+ -pyrophosphatase. *J. Biol. Chem.* **286**, 11970–11976 (2011).
21. Gordon-Weeks, R., Steele, S. H. & Leigh, R. A. The role of magnesium, pyrophosphate, and their complexes as substrates and activators of the vacuolar H^+ -pumping inorganic pyrophosphatase: studies using ligand protection from covalent inhibitors. *Plant Physiol.* **111**, 195–202 (1996).
22. Samygina, V. R. *et al.* Reversible inhibition of *Escherichia coli* inorganic pyrophosphatase by fluoride: trapped catalytic intermediates in cryo-crystallographic studies. *J. Mol. Biol.* **366**, 1305–1317 (2007).
23. Tzeng, C. M. *et al.* Subunit structure of vacuolar proton-pyrophosphatase as determined by radiation inactivation. *Biochem. J.* **316**, 143–147 (1996).
24. Zhen, R. G., Kim, E. J. & Rea, P. A. Acidic residues necessary for pyrophosphate-energized pumping and inhibition of the vacuolar H^+ -pyrophosphatase by *N,N'*-dicyclohexylcarbodiimide. *J. Biol. Chem.* **272**, 22340–22348 (1997).
25. Pan, Y. J. *et al.* The transmembrane domain 6 of vacuolar H^+ -pyrophosphatase mediates protein targeting and proton transport. *Biochim. Biophys. Acta* **1807**, 59–67 (2011).
26. Luecke, H., Richter, H. T. & Lanyi, J. K. Proton transfer pathways in bacteriorhodopsin at 2.3 angstrom resolution. *Science* **280**, 1934–1937 (1998).
27. Van, R. C. *et al.* Role of transmembrane segment 5 of the plant vacuolar H^+ -pyrophosphatase. *Biochim. Biophys. Acta* **1709**, 84–94 (2005).
28. Buch-Pedersen, M. J., Pedersen, B. P., Veierskov, B., Nissen, P. & Palmgren, M. G. Protons and how they are transported by proton pumps. *Pflügers Arch.* **457**, 573–579 (2009).
29. Pedersen, B. P., Buch-Pedersen, M. J., Morth, J. P., Palmgren, M. G. & Nissen, P. Crystal structure of the plasma membrane proton pump. *Nature* **450**, 1111–1114 (2007).
30. Huang, Y. T. *et al.* Distance variations between active sites of H^+ -pyrophosphatase determined by fluorescence resonance energy transfer. *J. Biol. Chem.* **285**, 23655–23664 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. F. Tam, P. C. Huang and H. J. Kung for their critical reading of the manuscript and for useful comments. The X-ray diffraction data were collected from the in-house X-ray facility at National Tsing Hua University and from beamlines BL13B1/BL13C1 at the National Synchrotron Radiation Research Center, Taiwan, and BL44XU/BL12B2 at SPring-8, Japan. This work was supported by grants from the National Science Council of Taiwan (NSC 99-2311-B-007-007-MY3 to Y.-J.S.; NSC 100-2311-B-007-001-MY3 and NSC 100-2627-M-007-012 to R.-L.P.) and National Tsing Hua University, Taiwan (99N82416E1 to Y.-J.S.).

Author Contributions S.-M.L. isolated the VrH^+ -PPase, grew the crystals and determined the structure. J.-Y.T. assisted with structural determination and completed the structural refinement. C.-D.H. assisted with structural phase analysis. C.-L.C. performed the data collection and data processing. J.-Y.T. and M.-H.L. assisted with the data collection. Y.-T.H. and T.-H.L. assisted in protein isolation. All authors participated in discussions of the results and in preparing the manuscript. Y.-J.S. and R.-L.P. supervised the project and wrote the manuscript.

Author Information The coordinates and structure factors of VrH^+ -PPase are deposited in the Protein Data Bank under the accession code 4A01. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.-J.S. (yjsun@life.nthu.edu.tw) or R.-L.P. (rlpan@life.nthu.edu.tw).

METHODS

Cloning, expression and purification. *Vigna radiata* H^+ -translocating pyrophosphatase (VrH^+ -PPase) fused with a C-terminal His-tag was expressed and isolated from *Saccharomyces cerevisiae* strain BJ2168 in accordance with a previously described method, with minor modifications³¹. Mutant derivatives were generated from wild-type VrH^+ -PPase by a QuikChange site-directed mutagenesis method³², and the sequences were verified by DNA sequencing. The VrH^+ -PPase was expressed in the yeast cells, which were transformed with the construct pYVH6 containing VrH^+ -PPase cDNA. The yeast cells were harvested by centrifugation at 4,000g for 10 min after induction for 3 days. The cell lysates were centrifuged at 4,000g for 10 min, and the supernatant was ultracentrifuged at 84,000g for 1 h for the collection of microsomal membranes. The membranes were suspended in a solution containing 25 mM MOPS-KOH pH 7.0, 20% (w/v) glycerol, 4 mM $MgCl_2$, 400 mM KCl and 1 mM phenylmethylsulphonyl fluoride (PMSF). The membrane proteins were solubilized from microsomal membranes with the use of n-dodecyl β -D-maltoside at a detergent-to-protein ratio of 3:1 (w/w) for 1 h at 4 °C. Insolubilized debris was removed by ultracentrifugation at 84,000g for 1 h. The solubilized VrH^+ -PPase was purified with Ni^{2+} -nitrilotriacetate resin (Qiagen, Valencia) and eluted with a solution containing 25 mM MOPS-KOH pH 7.0, 20% (w/v) glycerol, 4 mM $MgCl_2$, 400 mM KCl, 0.15% (w/v) n-decyl β -D-maltoside (DM) and 250 mM imidazole. The purified VrH^+ -PPase was dialysed against 25 mM MOPS-KOH pH 7.0, 20% (w/v) glycerol, 4 mM $MgCl_2$, 0.15% (w/v) DM, 5 mM IDP, and then concentrated to 10 mg ml⁻¹ for the crystallization setup. Enzymatic activities of microsomal and purified VrH^+ -PPases were assayed in accordance with previous methods³¹.

Trypsinolysis analysis. Trypsinolysis analysis was performed as described in a previous study²⁰. The microsomal (30 μ g) and purified (5 μ g) proteins were incubated with L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin at a ratio of 30:1 (w/w) at 37 °C for 10 min. The proteolysis was stopped by the addition of 2% (w/w) SDS and 5 mM PMSF. The samples were subjected to western blotting analysis.

Crystallization and X-ray data collection. The hanging-drop vapour-diffusion method³³ was used to set up crystallization trials. Both VrH^+ -PPase (0.5 μ l) and reservoir (0.5 μ l) solutions were mixed and equilibrated against a reservoir solution (500 μ l) in Linbro plates. VrH^+ -PPase crystals appeared in 2 days in the reservoir containing 23% (w/v) PEG2000, 50 mM MES pH 6.5, 10% (w/v) glycerol and 250 mM $MgCl_2$ at 20 °C. The native crystal data were collected on the BL44XU beamline at the SPring-8 synchrotron radiation facility, Japan. The data were processed using HKL2000 (ref. 34). VrH^+ -PPase crystals belong to the monoclinic space group C2 with the unit-cell parameters $a = 220.8$ Å, $b = 88.8$ Å, $c = 160.2$ Å and $\beta = 125.9^\circ$. The Matthews's coefficient³⁵ was calculated to be 3.97 Å³ Da⁻¹, corresponding to a solvent content of 69% with two subunits per asymmetric unit. For phase determination, three VrH^+ -PPase heavy-atom derivatives were obtained from Ta₆Br₁₂ (2 mM), OsO₄ (1 mM) and orange-Pt (1 mM). The anomalous data

were collected on the BL13B1 and BL13C1 beamlines at the National Synchrotron Radiation Research Center, Taiwan. The data statistics are summarized in Supplementary Table 1.

Structural determination and refinement. The structural phase of VrH^+ -PPase was obtained with a tantalum derivative by MAD and from OsO₄ and Pt-organic derivatives by MIRAS (Supplementary Table 1). The heavy-atom sites were located with SOLVE³⁶. The final phase combination was calculated by AutoSol using PHENIX^{37,38}, resulting in distinguishable protein and solvent regions. The preliminary structural model was automatically built by AutoBuild^{37,38}, and the entire model was completed manually with COOT³⁹. The refinement of the structure was performed with REFMAC5 (ref. 40). PROCHECK⁴¹ was used to evaluate the stereochemistry and to assign the secondary structural elements. The structural model was refined to an *R*-factor of 16.80% and an *R*_{free} of 20.31% at 2.35 Å. The refinement statistics are summarized in Supplementary Table 1. All the figures were generated using PYMOL⁴². The solvent-accessible surface area was calculated with the program HOLLOW⁴³ using a 1.4-Å probe radius.

- Hsu, S. H. *et al.* Purification, characterization, and spectral analyses of histidine-tagged vacuolar H^+ -pyrophosphatase expressed in yeast. *Bot. Stud. (Taipei, Taiwan)* **50**, 291–301 (2009).
- Kirsch, R. D. & Joly, E. An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. *Nucleic Acids Res.* **26**, 1848–1850 (1998).
- McPherson, A. Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **189**, 1–23 (1990).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
- Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849–861 (1999).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
- Vagin, A. A. *et al.* REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–2195 (2004).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
- DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific) (2002).
- Ho, B. K. & Gruswitz, F. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct. Biol.* **8**, 49 (2008).

The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria

Gene-Wei Li¹, Eugene Oh¹ & Jonathan S. Weissman¹

Protein synthesis by ribosomes takes place on a linear substrate but at non-uniform speeds. Transient pausing of ribosomes can affect a variety of co-translational processes, including protein targeting and folding¹. These pauses are influenced by the sequence of the messenger RNA². Thus, redundancy in the genetic code allows the same protein to be translated at different rates. However, our knowledge of both the position and the mechanism of translational pausing *in vivo* is highly limited. Here we present a genome-wide analysis of translational pausing in bacteria by ribosome profiling—deep sequencing of ribosome-protected mRNA fragments^{3–5}. This approach enables the high-resolution measurement of ribosome density profiles along most transcripts at unperturbed, endogenous expression levels. Unexpectedly, we found that codons decoded by rare transfer RNAs do not lead to slow translation under nutrient-rich conditions. Instead, Shine–Dalgarno-(SD)⁶-like features within coding sequences cause pervasive translational pausing. Using an orthogonal ribosome^{7,8} possessing an altered anti-SD sequence, we show that pausing is due to hybridization between the mRNA and 16S ribosomal RNA of the translating ribosome. In protein-coding sequences, internal SD sequences are disfavoured, which leads to biased usage, avoiding codons and codon pairs that resemble canonical SD sites. Our results indicate that internal SD-like sequences are a major determinant of translation rates and a global driving force for the coding of bacterial genomes.

Our current understanding of sequence-dependent translation rates *in vivo* derives largely from pioneering work begun in the 1980s^{9–13}. These studies, which measured protein synthesis times using pulse labelling, established that different mRNAs could be translated with different elongation rates. In particular, messages decoded by less abundant tRNAs were translated slowly, although this effect was exaggerated by the overexpression of mRNA, which can lead to the depletion of available tRNAs¹⁰. Even with fixed tRNA usage, different synonymously coded mRNAs were translated at different rates¹³. This result, together with the observation of biased occurrence of adjacent codon pairs¹⁴, suggested that tRNA abundance is not the only determinant of elongation rates. Further investigations addressing what determines the rate of translation *in vivo*, however, have been hampered by the limited temporal and positional resolution of existing techniques.

To provide a high-resolution view of local translation rates, we used the recently developed ribosome profiling strategy^{3–5} to map ribosome occupancy along each mRNA (Supplementary Fig. 1). We focused on two distantly related bacterial species, the Gram-negative bacterium *Escherichia coli* and the Gram-positive bacterium *Bacillus subtilis*. To preserve the state of translation, cells were flash-frozen in liquid nitrogen after the rapid filtration of exponential-phase cultures. Ribosome-protected footprints were generated through nuclease treatment of cell extract in the presence of inhibitors of translation elongation (see Methods). These steps ensured that most ribosomes were polysome-associated after lysis and stayed assembled as 70S particles during digestion (Supplementary Fig. 2). After deep sequencing, 2,257 genes in *E. coli* and 1,580 genes in *B. subtilis* had an average coverage of at least ten sequencing reads per codon. The observed variability of

ribosome footprint profiles across individual genes was highly reproducible ($r = 0.99$ between biological replicates; Supplementary Fig. 3).

Several observations argued that ribosome transit time is proportional to the occupancy at each position. First, we observed negligible internal initiation and early termination associated with ribosome pause sites (Supplementary Fig. 4). Second, ribosomes remained intact during footprinting, which enabled the large majority of ribosome-protected fragments to be captured (Supplementary Fig. 2). Third, the variability introduced during the conversion of RNA fragments into a sequenceable DNA library contributed minimally to our measures of variability in ribosome occupancy (Supplementary Fig. 5).

With our genome-wide view of local translation rates, we confirmed established examples of peptide-mediated stalling at transcripts *secM*¹⁵ and *tnaC*¹⁶ in *E. coli* and *mifM*¹⁷ in *B. subtilis* (Fig. 1a and Supplementary Fig. 6). Strikingly, in addition to these known stalling sites, the observed ribosome occupancy was highly variable across coding regions, as illustrated for *secA* in Fig. 1a. We found that ribosome density often surpasses by more than tenfold the mean density, and the vast majority of these translational pauses are uncharacterized.

We first sought to determine whether the identity of the codon being decoded could account for the differences in local translation rates, by examining the average ribosome occupancy for each of the 61 codons in the ribosomal A-site. Surprisingly, there was little correlation between the average occupancy of a codon and existing measurements of the abundance of corresponding tRNAs¹⁸ (Fig. 1b, c and Supplementary Fig. 7). Most notably, the six serine codons had the highest ribosome occupancy for *E. coli* cultured in Luria broth (Fig. 1b). Because serine is the first amino acid to be catabolized by *E. coli* when sugar is absent^{19,20}, we reasoned that the increased ribosome occupancy might be due to limited serine supply. Indeed, serine-associated pauses were greatly decreased in glucose-supplemented MOPS medium (Fig. 1c). The increase in serine codon occupancy when glucose becomes limiting confirmed our ability to capture translation rates at each codon. However, the identity of the A-site codon, which had less than a twofold effect on ribosome occupancy (Fig. 1c), could not account for the large variability in ribosome density along messages.

What, then, are the sequence features that cause slow translation? Without a priori knowledge about where such features would be located relative to the ribosomal A-site, we calculated the cross-correlation function between intragenic ribosome occupancy profiles and the presence of a given trinucleotide sequence on the mRNA independently of reading frames. Strong correlation was observed for six trinucleotide sequences (Fig. 1d) that resembled features found in Shine–Dalgarno (SD)⁶ sequences. The highest correlation occurred when the SD-like feature was 8–11 nucleotides upstream of the position occupied by the ribosomal A-site. This spacing coincides with the optimal spacing for ribosome binding at start codons²¹. However, unlike canonical SD sites, which enable initiation of translation, the observed pauses were associated with SD-like features within the body of coding regions. The accumulation of ribosomes at internal SD-like sequences was observed across two divergent phyla of bacteria (Fig. 2a), suggesting that the phenomenon occurs generally in bacteria.

¹Department of Cellular and Molecular Pharmacology, Howard Hughes Medical Institute, University of California, San Francisco, California 94158, USA.

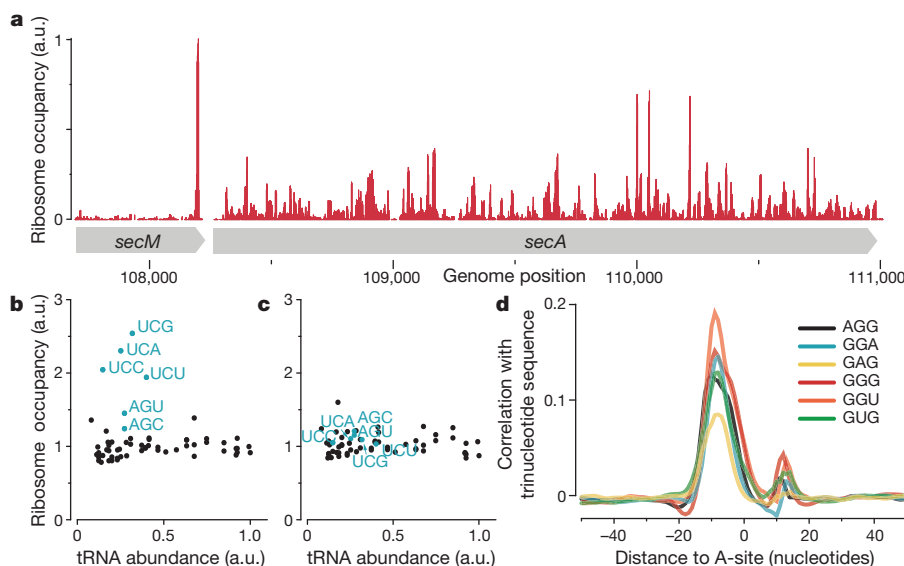


Figure 1 | Analysis of translational pausing using ribosome profiling in bacteria. **a**, Validation of the ribosome stalling site in the *secM* mRNA. **b**, **c**, Average ribosome occupancy of each codon relative to their respective tRNA abundances for *E. coli*. **b**, For growth in Luria broth, elevated occupancy at

serine codons (blue) probably reflects preferential depletion of this amino acid. **c**, In glucose-rich medium, the ribosome occupancy is independent of tRNA abundance. **d**, Plot of cross-correlation function between ribosome occupancy profiles and the presence of the indicated trinucleotide sequences for *E. coli*.

The same correlation was not observed for the budding yeast *Saccharomyces cerevisiae*, whose ribosomes, like those of other eukaryotes, do not contain an anti-SD (aSD) site.

As predicted by a model in which the interaction between mRNA and the aSD site of the 16S rRNA drives pausing, the predicted hybridization free energy of a hexanucleotide to the aSD sequence

was a strong indicator of its average downstream ribosome occupancy (Fig. 2b). Furthermore, there was a clear correspondence on individual transcripts between SD-like sequences and pauses. For example, Fig. 2c shows that in *ompF*, individual SD-like features are associated with elevated ribosome occupancy 8–11 nucleotides downstream. Moreover, a synonymous mutation (GGUGGU to GGCGGC) that decreased the affinity for the aSD site led to reduced ribosome occupancy specifically at the mutated sequence, suggesting a causal relationship between the SD-like feature and the excess ribosome density.

We next sought to evaluate directly whether the excess footprint density seen at internal SD-like sequences was due to pausing of elongating ribosomes rather than attempted internal initiation, driven by SD–aSD interactions (Fig. 3a). To distinguish between these possibilities, we used a previously described orthogonal ribosome (O-ribosome) system in which a mutant form of the 16S rRNA with an altered aSD site is expressed together with wild-type 16S rRNA⁸. O-ribosomes containing the mutant 16S RNA will only translate a target mRNA that has the corresponding orthogonal SD (O-SD) sequence before its start codon. Conversely, a message whose translation is driven by the O-SD sequence will only be translated by O-ribosomes, and not by wild-type ribosomes. This system thus allows one to determine the source of regions of excess ribosome footprints, because elongating O-ribosomes would pause at internal O-SD sequences, whereas attempted internal initiation would still occur at SD sequences as a result of the cellular pool of wild-type ribosomes.

We compared the ribosome occupancy profiles of a *lacZ* message that was translated by either O-ribosomes or wild-type ribosomes. The occupancy profile of the *lacZ* message exclusively translated by O-ribosomes was correlated with O-SD-like features, and not with SD-like features (Fig. 3c). This is in marked contrast with the same *lacZ* sequence translated by wild-type ribosomes (Fig. 3b). As an internal control in O-ribosome-expressing cells, all other genes, which were translated by wild-type ribosomes, still maintained SD-correlated ribosome occupancy profiles (Fig. 3c). These observations established that elongating ribosomes pause during protein synthesis and that hybridization between the aSD site in the elongating ribosome and internal SD-like sequences gives rise to these pauses.

Global analysis of pause sites revealed that internal SD-like sequences are the dominant feature controlling translational pausing: about 70% of the strong pauses (that is, those that have ribosome

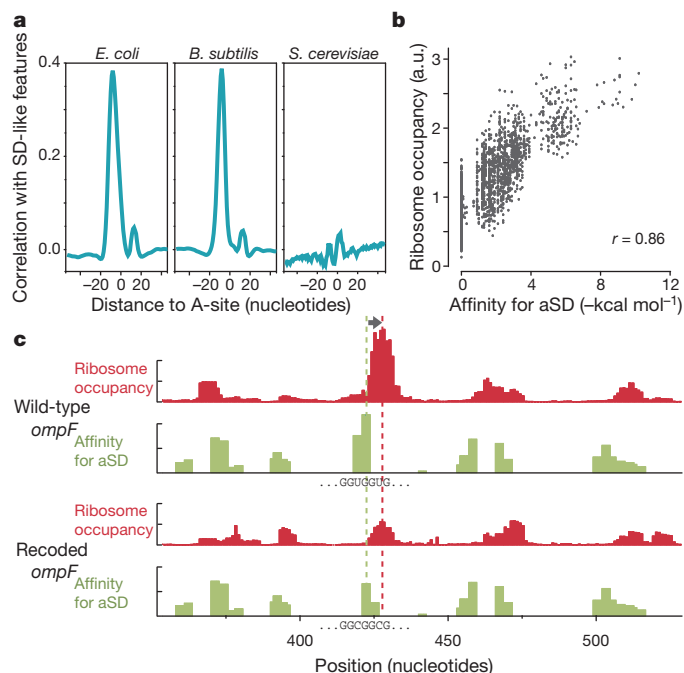


Figure 2 | Relationship between ribosome pausing and internal Shine-Dalgarno sequences. **a**, Plot of correlation between ribosome occupancy and SD-like features for *E. coli*, *B. subtilis* and *S. cerevisiae*. **b**, Plot of the average ribosome occupancy of hexanucleotide sequences relative to their affinity for the anti-Shine–Dalgarno sequence. **c**, Reprogrammed pausing by recoding the *ompF* mRNA. Ribosome occupancy (red) increases when the A-site is 8–11 nucleotides downstream (arrow) of SD-like features (green). Synonymous mutations replacing the SD-like sequence (GGUGGU) in wild-type *ompF* (top) with a sequence (GGCGGC) with lower affinity for the aSD site (bottom) caused a corresponding decrease in ribosome pausing.

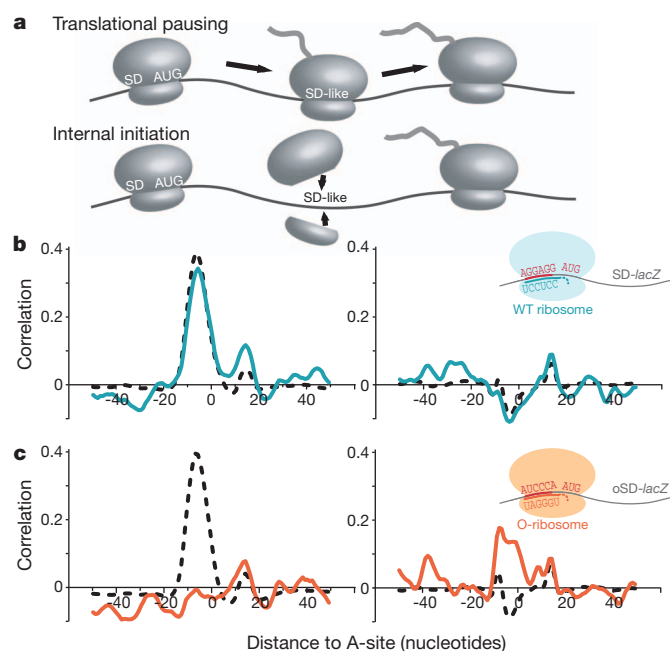


Figure 3 | Pausing of elongating ribosomes due to SD–aSD interaction. **a**, Two models could account for the excess ribosome density at internal SD-like sequences. **b**, Ribosome occupancy of *lacZ* mRNA translated by wild-type ribosome. Like other genes translated by the wild-type ribosome, the ribosome occupancy pattern along *lacZ* is correlated with the presence of SD-like sequences (left), not with the O-SD sequence (right). Cyan, *lacZ*; black, all other genes. **c**, Ribosome occupancy of *lacZ* mRNA translated by orthogonal ribosome (O-ribosome). Unlike other genes in the same cells, the specialized O-SD *lacZ* has ribosome pausing at internal O-SD-like sequences (right), not at SD-like sites (left). Orange, *lacZ*; black, all other genes.

occupancies more than tenfold over the mean) are associated with SD sites (Supplementary Fig. 8). Although the interaction between internal SD sequences in a message and elongating ribosomes has been documented in specialized cases, including promoting frame-shifting *in vivo*^{22,23} and ribosome stalling in single-molecule experiments *in vitro*²⁴, there was little indication that internal SD-like sequences are the major determinant of elongation rate during translation.

Because translational pausing limits the amount of free ribosomes available for initiating protein synthesis, widespread internal SD-like sequences could decrease bacterial growth rates. Accordingly, we found that strong SD-like sequences are generally avoided in the coding region of *E. coli* genes: hexamer sequences that strongly bind aSD sites are universally rare, whereas low-affinity hexamers have variable rates of occurrence (Fig. 4a). Consistent with translational

pausing being the driving force for this bias, depletion of SD-like sequences was observed only in protein-coding genes, and not in genes encoding rRNA or tRNA (Supplementary Fig. 9). The selection against SD-like features in turn impacts both synonymous codon choice and codon-pair choice. At the codon level, SD-like codons GAG, AGG and GGG are all minor codons in *E. coli* and *B. subtilis*. The evolutionary origin of codon selection is often attributed to differences in tRNA abundance^{2,25} because its level is correlated with codon usage¹⁸. Instead, we propose that SD-like codons are disfavoured as a result of their interactions with rRNA, and that tRNA expression levels followed codon adaptation.

At the codon-pair level, we can now account for the selection against two consecutive codons that resemble SD sequences. This is illustrated for Gly-Gly pairs, which are coded by GGNNGN sequences (Fig. 4b). The most abundant Gly-Gly coding sequence, GGCGGC, has the lowest affinity for the aSD sequence, whereas Gly-Gly coding sequences that strongly resemble SD sites, including GGAGGU, which perfectly complements the aSD site, rarely appear. This under-representation holds even after correcting for the usage of individual codons (Fig. 4b, colour coding); for example, GGAGGU is considerably less common than GGUGGA. Other amino-acid pairs that can be coded with strong SD sites also showed the same bias (Supplementary Fig. 10). The preference in codon pairs stems from the sequence identity and not codon identity, because the same trend is seen in hexamers that are not aligned to codon pairs (Supplementary Fig. 11). Although not every bias in codon-pair usage can be explained here, the disadvantage associated with SD-induced translational pausing offers a clear mechanistic view of why certain codon pairs are avoided.

Despite the selection against internal SD-like sequences, they remain a major driving force for translational pausing. In addition, we found similar pausing patterns between conserved genes in *E. coli* and *B. subtilis* (Fig. 4c). For an mRNA encoding a specific protein, it may not be possible to fully eliminate sequences with affinity for the aSD site without changing the peptide sequence. For example, in the case of Gly-Gly, even the GGCGGC pair has substantial affinity for the aSD site. The optimization for translation rate therefore cannot be achieved only at the level of mRNA coding: it is also constrained by the requirement to make a functional peptide sequence.

The observation that the ability of elongating ribosomes to interact with SD-like sequences is highly conserved suggests that this mechanism of pausing is exploited for functional purposes. Indeed, a highly conserved internal SD site exists in the gene encoding peptide chain release factor 2 (RF2)²⁶. This sequence has an important function in promoting a translational frameshift to enable its expression. In addition, pausing at internal SD-like sites could modulate the co-translational folding of the nascent peptide chain (Supplementary Fig. 12). Finally, given the coupling between transcription and translation in bacteria^{27,28}, pausing

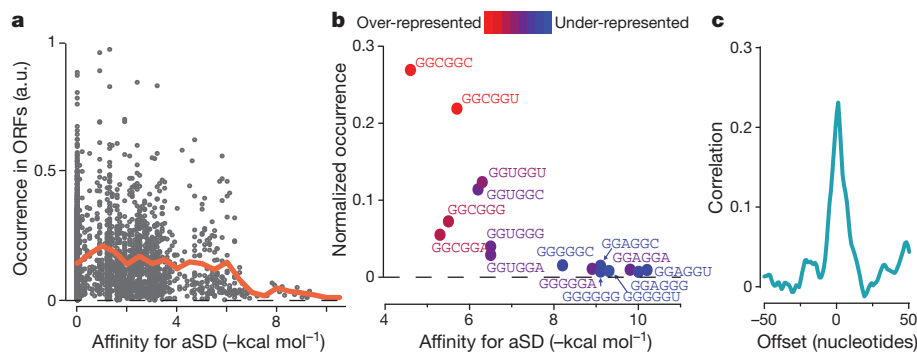


Figure 4 | Selection against SD-like sequences and the constraint on protein coding. **a**, Rate of occurrence of hexanucleotide sequences in *E. coli* messages relative to their predicted affinity for the aSD site. The orange line shows the average occurrence within a bin size of 0.5 kcal mol^{–1}. **b**, Occurrence of codon-pairs for Gly-Gly residues relative to their predicted affinity for the aSD site. The colour coding represents the enrichment in occurrence of codon pairs after correcting for the usage of single codons. **c**, Cross-correlation function of ribosome occupancy profiles between conserved genes in *E. coli* and *B. subtilis*. Zero offset means that the two sequences are aligned at each amino-acid residue.

The colour coding represents the enrichment in occurrence of codon pairs after correcting for the usage of single codons. **c**, Cross-correlation function of ribosome occupancy profiles between conserved genes in *E. coli* and *B. subtilis*. Zero offset means that the two sequences are aligned at each amino-acid residue.

at SD sites could be exploited for transcriptional regulation. We observed internal SD sites and pausing near the stop codon of transcription attenuation leader peptides²⁹, including *trpL* and *thrL* (Supplementary Fig. 13). In contrast to ribosome stalling at regulatory codons during starvation, slow translation near the stop codon could protect alternative structural mRNA elements to prevent the formation of anti-termination stem-loops, thereby ensuring proper transcription termination³⁰. Our approach and the genome-wide data lay the groundwork for further gene-specific functional studies of translational pausing.

From a more practical perspective, ribosome pausing at internal SD sites presents both a challenge and an opportunity for heterologous protein expression in bacteria. Overexpression of eukaryotic proteins with strong internal SD sites would sequester ribosomes and compromise protein yield. Internal SD sequences could be reduced by recoding the gene, which has not been considered in conventional strategies of simple codon optimization or overexpression of rare tRNAs. Conversely, recoding can introduce internal SD sites if pausing is required for co-translational processing. Positioning of internal SD sites therefore adds another dimension to the optimization of heterologous protein expression.

METHODS SUMMARY

E. coli MG1655 and *B. subtilis* 168 were used as wild-type strains. *E. coli* BJW9 has synonymous substitutions at G141 and G142 in the *ompF* gene. The orthogonal ribosome experiment was performed in *E. coli* BW25113 with two plasmids: pSC101-G9, expressing orthogonal 16S rRNA, and pJW1422, expressing O-SD-*lacZ* mRNA. pSC101-G9 was a gift from J. Chin⁸. pJW1422 has *lacZ* driven from a *tacII* promoter and an O-ribosome binding site 5'-AUCCCA-3'. Luria broth was used for *B. subtilis* culture. Cell cultures were harvested at a D_{600} of 0.3–0.4. Flash-freezing and ribosome footprinting was described previously⁵. 5'-Guanylyl imidodiphosphate (3 mM) was added to the lysate before thawing and during footprinting to prevent translation after lysis. Conversion of mRNA footprints to a complementary DNA library was described previously^{4,5}. Deep sequencing was performed on an Illumina HiSeq 2000 system, and the results were aligned to reference genomes using Bowtie v. 0.12.0. The cross-correlation function is defined as

$$C_i = \frac{\langle x_{j+i} y_j \rangle - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

for the series $X = x_1, x_2, \dots, x_N$ and $Y = y_1, y_2, \dots, y_N$, where μ_X and σ_X are the average and the standard deviation of series X , respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 November 2011; accepted 16 February 2012.

Published online 28 March 2012.

- Kramer, G., Boehringer, D., Ban, N. & Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature Struct. Mol. Biol.* **16**, 589–597 (2009).
- Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Rev. Genet.* **12**, 32–42 (2011).
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
- Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789–802 (2011).
- Oh, E. *et al.* Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in vivo*. *Cell* **147**, 1295–1308 (2011).
- Shine, J. & Dalgarno, L. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl Acad. Sci. USA* **71**, 1342–1346 (1974).
- Hui, A. & de Boer, H. A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **84**, 4762–4766 (1987).

- Rackham, O. & Chin, J. W. A network of orthogonal ribosome-mRNA pairs. *Nature Chem. Biol.* **1**, 159–166 (2005).
- Varenne, S., Buc, J., Llobes, R. & Lazdunski, C. Translation is a non-uniform process. Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *J. Mol. Biol.* **180**, 549–576 (1984).
- Pedersen, S. *Escherichia coli* ribosomes translate *in vivo* with variable rate. *EMBO J.* **3**, 2895–2898 (1984).
- Sorensen, M. A., Kurland, C. G. & Pedersen, S. Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* **207**, 365–377 (1989).
- Andersson, S. G. & Kurland, C. G. Codon preferences in free-living microorganisms. *Microbiol. Rev.* **54**, 198–210 (1990).
- Sorensen, M. A. & Pedersen, S. Absolute *in vivo* translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J. Mol. Biol.* **222**, 265–280 (1991).
- Gutman, G. A. & Hatfield, G. W. Nonrandom utilization of codon pairs in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **86**, 3699–3703 (1989).
- Nakatogawa, H. & Ito, K. The ribosomal exit tunnel functions as a discriminating gate. *Cell* **108**, 629–636 (2002).
- Gong, F. & Yanofsky, C. Instruction of translating ribosome by nascent peptide. *Science* **297**, 1864–1867 (2002).
- Chiba, S. *et al.* Recruitment of a species-specific translational arrest module to monitor different cellular processes. *Proc. Natl Acad. Sci. USA* **108**, 6073–6078 (2011).
- Dong, H., Nilsson, L. & Kurland, C. G. Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* **260**, 649–663 (1996).
- Pruss, B. M., Nelms, J. M., Park, C. & Wolfe, A. J. Mutations in NADH:ubiquinone oxidoreductase of *Escherichia coli* affect growth on mixed amino acids. *J. Bacteriol.* **176**, 2143–2150 (1994).
- Sezonov, G., Joseleau-Petit, D. & D'Ari, R. *Escherichia coli* physiology in Luria-Bertani broth. *J. Bacteriol.* **189**, 8746–8749 (2007).
- Chen, H., Bjerknes, M., Kumar, R. & Jay, E. Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res.* **22**, 4953–4957 (1994).
- Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F. & Gesteland, R. F. Reading frame switch caused by base-pair formation between the 3' end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*. *EMBO J.* **7**, 1503–1507 (1988).
- Larsen, B., Wills, N. M., Gesteland, R. F. & Atkins, J. F. rRNA-mRNA base pairing stimulates a programmed –1 ribosomal frameshift. *J. Bacteriol.* **176**, 6842–6851 (1994).
- Wen, J. D. *et al.* Following translation by single ribosomes one codon at a time. *Nature* **452**, 598–603 (2008).
- Ikemura, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* **151**, 389–409 (1981).
- Baranov, P. V., Gesteland, R. F. & Atkins, J. F. Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3**, 373–377 (2002).
- Burmann, B. M. *et al.* A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).
- Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010).
- Kolter, R. & Yanofsky, C. Attenuation in amino acid biosynthetic operons. *Annu. Rev. Genet.* **16**, 113–134 (1982).
- Elf, J. & Ehrenberg, M. What makes ribosome-mediated transcriptional attenuation sensitive to amino acid limitation? *PLOS Comput. Biol.* **1**, e2 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Reuman, D. Burkhardt, C. Jan, C. Gross, J. Elf and members of the Weissman laboratory for discussions; J. Dunn for ribosome profiling data on *S. cerevisiae*; C. Chu for help with sequencing; and J. Chin for orthogonal ribosome reagents and advice. This research was supported by the Helen Hay Whitney Foundation (to G.W.L.) and by the Howard Hughes Medical Institute (to J.S.W.).

Author Contributions G.W.L. and J.S.W. designed the experiments. G.W.L. performed experiments and analysed the data. E.O. provided technical support and preliminary data. G.W.L. and J.S.W. wrote the manuscript.

Author Information The footprint sequencing data are deposited in the Gene Expression Omnibus (GEO) under accession number GSE35641. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.S.W. (weissman@cmp.ucsf.edu).

METHODS

Strains, plasmids and growth conditions. *E. coli* K-12 MG1655 and *B. subtilis* 168 were used as wild-type strains. *E. coli* BJW9 with recoded *ompF* (GGT to GGC synonymous substitutions at G141 and G142) at the endogenous locus was constructed by lambda-Red recombination³¹ in MG1655. The orthogonal ribosome experiment was performed in *E. coli* BW25113, which is a K-12-derived strain with a *lacZ* deletion³¹.

Plasmid pSC101-G9 (gift from J. Chin), expresses orthogonal 16S rRNA from an intact *rrnB* operon except that the 3' end of *rrsB*, which codes for the 16S rRNA, was changed from 5'-CCTCCTTA-3' to 5'-TGGGATTA-3' (ref. 8). Plasmid pJW1422 harbours the *lacZ* gene with a *tacII* promoter. The ribosome-binding site of the *lacZ* mRNA is replaced with 5'-AUCCCA-3', thus allowing initiation of translation by orthogonal ribosomes.

Unless otherwise noted, *E. coli* strains were grown in MOPS medium supplemented with 0.2% glucose, 20 amino acids, vitamins, bases and micronutrients as described³² (Teknova). *B. subtilis* was grown in Luria broth (BD Difco). For the strain containing pSC101-G9 and pJW1422, the medium was supplemented with 25 µg ml⁻¹ kanamycin and 15 µg ml⁻¹ tetracycline. For experiments with *E. coli*, an overnight liquid culture was diluted 1:400 into fresh medium. For experiments with *B. subtilis*, an overnight culture on a Luria broth plate was washed and diluted to a D_{600} of 0.00125 in Luria broth. Cell cultures were grown at 37 °C until D_{600} reached 0.3–0.4.

Ribosome profiling. The protocol for bacterial ribosome profiling with flash freezing was described in ref. 5. Cell culture (200 ml) was rapidly filtered by passing through a prewarmed nitrocellulose filter with a 200-nm pore size. Cell pellet was flash-frozen in liquid nitrogen and combined with 650 µl of frozen lysis buffer (10 mM MgCl₂, 100 mM NH₄Cl, 20 mM Tris-HCl pH 8.0, 0.1% Nonidet P40, 0.4% Triton X-100, 100 U ml⁻¹ DNase I (Roche), 1 mM chloramphenicol, 3 mM 5'-guanylyl imidodiphosphate (GMPPNP)). Addition of GMPPNP together with chloramphenicol inhibits translation after lysis. Cells were pulverized in 10-ml canisters prechilled in liquid nitrogen. Lysate containing 0.5 mg of RNA was digested for 1 h with 750 U of micrococcal nuclease (Roche) at 25 °C. The ribosome-protected fragments were isolated using a sucrose gradient and phenol extraction. The footprints were ligated to a 5'-adenylated and 3'-end-blocked DNA oligonucleotide (5rApp/CTGTAGGCACCATCAAT/3ddc; Integrated DNA Technologies)^{4,5}. Unless otherwise noted, the ligation was performed with truncated T4 RNA ligase 2 (New England Biolabs) as described previously^{4,5}. To remove lot-to-lot difference in the activity from the commercial source, we have recently switched to recombinantly expressed truncated T4 RNA ligase 2 K227Q produced in our laboratory. We used this ligase to generate a library for the high-coverage data set for *E. coli*. The 3'-ligated RNA fragments were converted to sequenceable DNA library by using reverse transcription, circularization and PCR amplification as described previously^{4,5}.

Sequencing was performed on an Illumina HiSeq 2000 system. Sequence alignment with Bowtie v. 0.12.0 mapped the footprint data to the reference genomes NC_000913.fna (*E. coli*) or NC_000964.fna (*B. subtilis*) obtained from the NCBI Reference Sequence Bank. The data from BJW9 were aligned to a reference modified from NC_000913.fna. The footprint reads varied between 25 and 42 nucleotides in length, mostly as a result of the specificity of micrococcal nuclease. In contrast to eukaryotic systems, in which the 5' end of the footprint is sufficient to carry the positional information^{3,4}, here we distribute the positional information into several nucleotides in the centre of the footprint⁵. For each footprint read, the centre residues that were at least 12 nucleotides from either end were given the same score, which was weighted by the length of the fragment.

To assign the A-site position to the centre of ribosome footprints, we made use of the ribosome density at two independent sets of well-defined pause sites. The first set consisted of pausing at stop codons⁵, where the ribosomal A-site was aligned to stop codons before binding of release factors. The second set consisted of peptide-mediated ribosome stalling sites, where the A-site codons had been identified. These two alignments were consistent with each other. In addition, the pausing at serine codons at the A-site during starvation confirmed the position assignment of ribosome footprints.

mRNA sequencing. Total RNA was phenol extracted from the same lysate that was used for ribosome footprinting⁵. Ribosomal RNA and small RNA were removed from the total RNA with MICROBExpress and MEGAclear (Ambion), respectively. mRNA was randomly fragmented as described⁴. The fragmented mRNA sample was converted to a complementary DNA library with the same strategy as for ribosome footprints, and was described previously⁵.

Data analysis. Data analysis was performed with scripts written for Python 2.6.6. Global pausing analyses were based on 2,257 genes (*E. coli*) and 1,580 genes (*B. subtilis*) with an average coverage of at least ten sequencing reads per codon in the ribosome profiling data set. In addition, analyses on 997 genes in *E. coli* and 1,189 genes in *B. subtilis* with an average coverage of between one and ten sequencing reads per codon showed qualitatively consistent results. For *E. coli*, *tufA* and *tufB* genes

were not included in the analysis because of their sequence homology with each other. Genes with known frame-shifting sites (*prfB* and *dnaX*) were not included in codon-specific analyses. On gene-specific analyses, the coverage was at least 30 sequencing reads per codon in each case.

To focus on the kinetics of translation elongation, the analysis was performed on the basis of ribosome occupancy profiles within protein-coding genes, excluding the first ten codons and the last ten codons. To calculate the average ribosome occupancy associated with each codon at the A-site, the ribosome occupancy profile of each gene was normalized by the mean occupancy of the gene, and the normalized occupancy for each codon was averaged across all genes. Similarly, the average ribosome occupancy for each hexanucleotide at the SD position was calculated by averaging normalized occupancy at between 7 and 12 nucleotides downstream of the hexanucleotide sequence. For each codon, the corresponding tRNA abundance plotted in Fig. 1 and Supplementary Fig. 7 was the sum of the expression levels of the cognate tRNA species measured in refs 18, 33.

To identify dominating sequence features either upstream or downstream of the pausing sites, we slid the ribosome occupancy profile ($X = x_1, x_2, \dots, x_N$) along the coding sequence and, at every offset position i , calculated the correlation with the presence of a given sequence ($Y = y_1, y_2, \dots, y_N$). In mathematical terms it is given by the normalized cross-correlation function (C_i):

$$C_i = \frac{\langle x_{j+i}y_j \rangle - \mu_X\mu_Y}{\sigma_X\sigma_Y}$$

where μ_X and μ_Y are the average of the series X and Y , respectively. σ_X and σ_Y are the standard deviations of the series X and Y , respectively. $\langle x_{j+i}y_j \rangle$ is the expectation value of $x_{j+i}y_j$ for all possible values of j . We used Python to calculate $\sum x_{j+i}y_j$ using the 'correlate' function in the 'same' mode in the numpy package. The expectation value is obtained by dividing the summation by $N - |i|$. For each gene with more than ten sequencing reads per codon and more than 160 base pairs long, we calculated the normalized cross-correlation function. The average over these cross-correlation functions is presented in this paper.

Hybridization free-energy prediction. The hybridization free energy between mRNA and the aSD site was predicted with the RNAsubopt program in the Vienna RNA package³⁴. The energy was predicted for 37 °C with a contribution from dangling ends. For each hexanucleotide sequence, the lowest possible hybridization free energy was assigned as its affinity for the aSD site. We used the eight-nucleotide sequence 5'-CACCUCU-3' as the aSD sequence. To calculate the cross-correlation function between ribosome occupancy profile and SD-like features (Fig. 2a), we built the aSD affinity profile for each mRNA by scanning the transcript in overlapping units of ten nucleotides and calculating the affinity of aSD to the middle eight nucleotides. The affinity was assigned to the eighth position in the ten-nucleotide window, which corresponds to U in the canonical SD sequence. The distance from the P-site to U in the canonical SD sequence is often defined as the aligned spacing²¹. Because we align ribosome footprints to the A-site, the distance reported here is three nucleotides longer than the aligned spacing.

Analysis of O-ribosome translated messages. Because a *lacZ* message whose translation is driven by O-SD is exclusively translated by O-ribosomes⁸, the translational pausing model outlined in Fig. 3a predicts that for the O-SD driven *lacZ* there will be both the appearance of new ribosome density peaks at internal sites that resemble the O-SD sequence and the disappearance of peaks at the SD-like sequences found when translation is driven by the wild-type SD sequence. This prediction is confirmed by our data in Fig. 3c: the ribosome occupancy profile of *lacZ* with O-SD-driven translation no longer shows a correlation with SD-like sequences, and instead is correlated with O-SD-like sequences. Moreover, because the endogenous messages are still translated solely by wild-type ribosomes even when the O-ribosome is present, the ribosome peaks in the endogenous messages are found at SD-like sequences, not at sequences that resemble the O-SD site, regardless of whether O-ribosomes are present. This is again confirmed by the data shown in Fig. 3c.

Conservation analysis. Conservation analysis of pausing patterns in *E. coli* and *B. subtilis* was performed in a set of 31 proteins from the curated alignment database AMPHORA³⁵. The nucleotide sequences and the ribosome density profiles were trimmed and concatenated on the basis of the protein alignment. The cross-correlation function between *E. coli* and *B. subtilis* ribosome occupancy was calculated for each gene, and then averaged over 31 genes to give the conservation of pausing patterns.

Occurrence of hexamers and codon pairs. The occurrence of hexamers and codon pairs was counted from annotated protein-coding genes available from the NCBI Reference Sequence Bank. Normalized occurrence (p_{ij}) was calculated by dividing the occurrence of a given codon pair (i and j) by the total occurrence of the corresponding amino-acid pair. The correction for the usage of single codons

was calculated by dividing the normalized occurrence of the codon pair (p_{ij}) by the frequency of the two individual codons (q_i and q_j); that is, enrichment = $p_{ij}/q_i q_j$. The frequency of individual codons was normalized to the occurrence of the corresponding amino acid.

Protein structure analysis. Protein secondary structure was predicted by the PSIPRED method³⁶, with the filtered reference database UniRef90 (ref. 37). Secondary structures were predicted for 271 proteins. Cross-correlation function between the structural assignment with either ribosome occupancy or SD-like features was calculated at the nucleotide level.

31. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
32. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974).
33. Kanaya, S., Yamada, Y., Kudo, Y. & Ikemura, T. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* **238**, 143–155 (1999).
34. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res.* **36** (suppl. 2), W70–W74 (2008).
35. Wu, M. & Eisen, J. A. A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* **9**, R151 (2008).
36. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
37. Li, W., Jaroszewski, L. & Godzik, A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**, 282–283 (2001).

Exploiting a natural conformational switch to engineer an interleukin-2 ‘superkine’

Aron M. Levin^{1*}, Darren L. Bates^{2*}, Aaron M. Ring^{2*}, Carsten Krieg^{3,4*}, Jack T. Lin⁵, Leon Su⁵, Ignacio Moraga², Miro E. Raeber^{3,4}, Gregory R. Bowman⁶, Paul Novick⁶, Vijay S. Pande⁶, C. Garrison Fathman⁵, Onur Boyman^{3,4} & K. Christopher Garcia^{1,2}

The immunostimulatory cytokine interleukin-2 (IL-2) is a growth factor for a wide range of leukocytes, including T cells and natural killer (NK) cells^{1–3}. Considerable effort has been invested in using IL-2 as a therapeutic agent for a variety of immune disorders ranging from AIDS to cancer. However, adverse effects have limited its use in the clinic. On activated T cells, IL-2 signals through a quaternary ‘high affinity’ receptor complex consisting of IL-2, IL-2R α (termed CD25), IL-2R β and IL-2R γ ^{4–8}. Naive T cells express only a low density of IL-2R β and IL-2R γ , and are therefore relatively insensitive to IL-2, but acquire sensitivity after CD25 expression, which captures the cytokine and presents it to IL-2R β and IL-2R γ . Here, using *in vitro* evolution, we eliminated the functional requirement of IL-2 for CD25 expression by engineering an IL-2 ‘superkine’ (also called super-2) with increased binding affinity for IL-2R β . Crystal structures of the IL-2 superkine in free and receptor-bound forms showed that the evolved mutations are principally in the core of the cytokine, and molecular dynamics simulations indicated that the evolved mutations stabilized IL-2, reducing the flexibility of a helix in the IL-2R β binding site, into an optimized receptor-binding conformation resembling that when bound to CD25. The evolved mutations in the IL-2 superkine recapitulated the functional role of CD25 by eliciting potent phosphorylation of STAT5 and vigorous proliferation of T cells irrespective of CD25 expression. Compared to IL-2, the IL-2 superkine induced superior expansion of cytotoxic T cells, leading to improved antitumour responses *in vivo*, and elicited proportionally less expansion of T regulatory cells and reduced pulmonary oedema. Collectively, we show that *in vitro* evolution has mimicked the functional role of CD25 in enhancing IL-2 potency and regulating target cell specificity, which has implications for immunotherapy.

To engineer a CD25-independent version of IL-2, we displayed human IL-2 on the surface of yeast as a conjugate to Aga2p, and verified proper receptor-binding properties with IL-2R β and IL-2R γ ectodomain tetramers that were carboxy-terminally biotinylated and coupled to phycoerythrin-conjugated streptavidin for use as a staining and sorting reagent^{9,10}. Yeast-displayed IL-2 bound to IL-2R γ in the presence of IL-2R β , recapitulating the cooperative assembly of the heterodimeric receptor complex as seen with soluble IL-2 (Fig. 1a and Supplementary Fig. 1). We proceeded to carry out two generations of *in vitro* evolution (Fig. 1b and Supplementary Fig. 2). Our first generation *in vitro* evolution strategy was to create an error-prone PCR library of the entire IL-2 coding sequence (Supplementary Fig. 2), which resulted in selection of a predominant IL-2 variant containing an L85V mutation (Fig. 1c and Supplementary Fig. 3).

From inspection of the wild-type (WT) IL-2 structure, we were surprised to find that position 85 was not a direct IL-2R β contact residue, but rather resided on the internal face of the IL-2 C-helix, within the hydrophobic core of the cytokine (Fig. 2a). Thus, we surmised

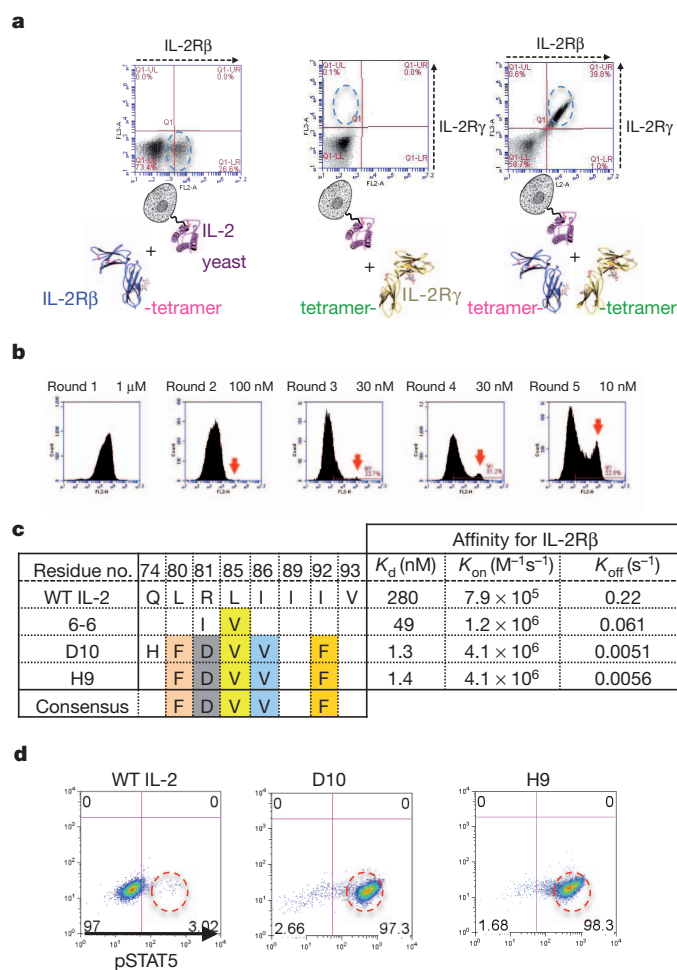


Figure 1 | *In vitro* evolution of human IL-2 variants with high affinity for IL-2R β . **a**, IL-2 displayed on yeast recapitulates cooperative receptor-binding activity. As measured by flow cytometry, IL-2 binds weakly to IL-2R β (left panel), undetectably to IL-2R γ (middle panel), and cooperatively forms the IL-2R β –IL-2R γ heterodimer (right panel). **b**, Enrichment of IL-2 variants on yeast by selection with progressively lower concentrations of IL-2R β . Red arrows indicate an emerging population of high-affinity IL-2R β binders (see also Supplementary Fig. 2). **c**, Sequences and affinities for IL-2R β of selected mutants from the first (mutant 6-6) and second (mutants D10 and H9) generation libraries (see Supplementary Fig. 3 for an extended table). **d**, On-yeast stimulation of YT-1 cells (human NK cell line) by wild-type (WT) IL-2-yeast and high-affinity variants (IL-2 superkines) (see also Supplementary Fig. 4).

¹Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA. ²Department of Molecular and Cellular Physiology, and Department of Structural Biology, Stanford University School of Medicine, Stanford, California 94305, USA. ³Laboratory of Applied Immunobiology, University of Zurich, Zurich CH-8006, Switzerland. ⁴Allergy Unit, Department of Dermatology, University Hospital Zurich, Zurich CH-8091, Switzerland. ⁵Stanford University School of Medicine, Department of Medicine, Division of Immunology and Rheumatology, Stanford, California 94305, USA. ⁶Department of Chemistry, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

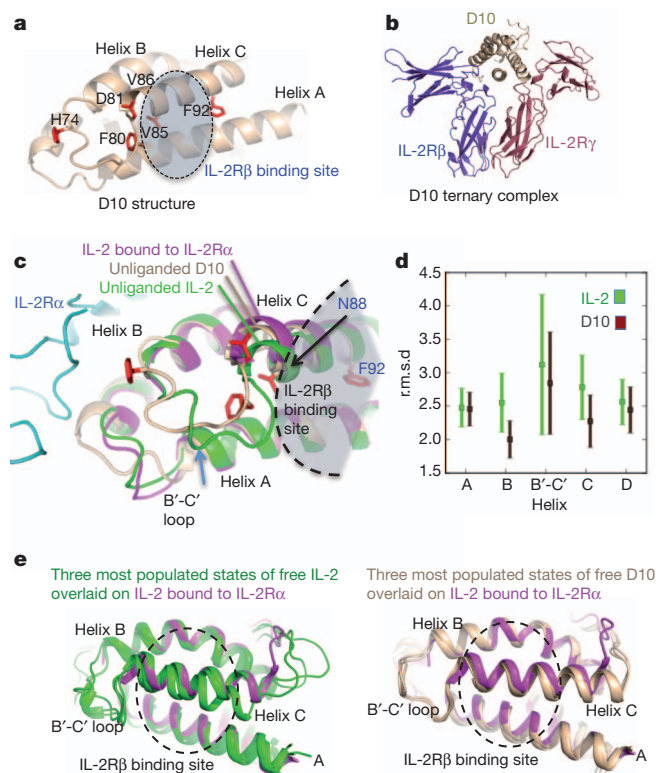


Figure 2 | Basis of affinity enhancement for IL-2R β from structural and molecular dynamics characterization of the D10 IL-2 superkine. **a**, Crystal structure of the D10 IL-2 superkine at 3.1 Å with mutated residues in red (see also Supplementary Table 1 and Supplementary Fig. 7a). **b**, D10 in complex with human IL-2R β and IL-2R γ preserves the wild-type receptor dimer geometry (see also Supplementary Fig. 7b). **c**, The unliganded D10 IL-2 superkine helix C (brown), moves towards its hydrophobic core compared to unliganded wild-type IL-2 (green, PDB ID 3INK). This helix C position is more similar to that of helix C in IL-2 bound to IL-2R α (purple, PDB ID 1Z92) (see also Supplementary Fig. 8). **d**, A 40-ns molecular dynamics simulation shows a reduction of the average r.m.s.d. for the B and C helices, and the B-C loop in D10 versus IL-2 (see also Supplementary Fig. 8c). Error bars represent the standard error of the r.m.s.d. **e**, Helix C in IL-2 (green, left panel) drifts during the molecular dynamics simulation more than the IL-2 superkine D10 (brown, right panel) when compared to IL-2 bound to IL-2R α (purple).

that L85V may affect the structure of helix C in a way that enhances binding to IL-2R β . Therefore, we carried out a second generation selection where we made a biased library that contained F/I/L/V at amino acids L80, L85, I86, I89, I92 and V93, which are contained within the hydrophobic core and linker region on helix C (Fig. 1b, c). To rapidly select the most active variants, we used the yeast-displayed cytokines themselves to stimulate STAT5 phosphorylation in the human NK cell line YT-1 by co-incubation at varying yeast:YT-1 cell ratios (Fig. 1d and Supplementary Fig. 4). Several clones stimulated substantially more STAT5 phosphorylation at lower yeast:cell ratios than yeast-displayed wild-type IL-2 (Fig. 1d and Supplementary Fig. 4). Sequencing of a selected panel of high-affinity IL-2 clones revealed a consensus set of mutations L80F/R81D/L85V/I86V/I92F (Fig. 1c and Supplementary Fig. 3).

We expressed recombinant forms of several first- and second-generation IL-2 clones to measure their binding affinities and kinetics for IL-2R β by surface plasmon resonance (SPR) (Fig. 1c and Supplementary Figs 3 and 5) and isothermal titration calorimetry (ITC) (Supplementary Fig. 6). By SPR, the affinity between IL-2 and IL-2R β was $K_d = 280$ nM. The IL-2 superkines, also called 'super-2s', clustered into low, medium and high affinity classes. The highest affinity mutants had K_d s of 1.2–1.7 nM (D10, H9). The affinity increases were

uniformly manifested in reductions in off-rates (Fig. 1c and Supplementary Figs 3 and 5).

To understand the structural consequences of the evolved mutations, we crystallized the D10 IL-2 superkine (Fig. 2a, Supplementary Fig. 7a and Supplementary Table 1). In the structure of D10 alone, five of the six mutations clustered on the B-C loop and within the C-helix core, in positions that did not contact IL-2R β . Notably, the B-C-helix linker region was ordered in the electron density map (Supplementary Figs 7a and 8a), compared to other IL-2 structures where this region is often disordered (Supplementary Fig. 8a). Collectively, the F80, V85 and V86 substitutions appeared to collapse into a hydrophobic cluster that stabilized the loop by pinning the C-helix into the core of the molecule. Only one of the five consensus mutations, I92F, was at a position that contacted IL-2R β in the receptor complex (Fig. 2a), but it was deeply inserted between the C and A helices, contributing only an additional 10 Å² of molecular surface buried by IL-2R β in the complex compared to Ile 92. We also determined a low-resolution (3.8 Å) structure of the D10 ternary receptor complex to assess whether the mutations have perturbed the IL-2R β /IL-2R γ receptor dimer geometry compared to the wild-type IL-2 complex (Fig. 2b and Supplementary Table 1). The overall IL-2R β /IL-2R γ heterodimeric architecture and mode of cytokine/IL-2R β contact in the D10 ternary complex were essentially identical to the previously reported IL-2 quaternary assembly (root mean squared deviation (r.m.s.d.) = 0.43 Å) (Supplementary Fig. 7b).

Previously, we found that the C-helix of IL-2 seems to undergo subtle repositioning upon binding to IL-2R α ¹¹ (Fig. 2c and Supplementary Fig. 8a). Inspection of three wild-type unliganded IL-2 structures revealed conformational variability in the C-helix position, consistent with higher crystallographic B-factors in this helix relative to the rest of the molecule (Supplementary Fig. 8b). We compared the structure of our D10 IL-2 superkine to that of an unliganded structure of IL-2, and IL-2 in the receptor complexes. We found that the C-helix in D10 was more similar to that seen in the two receptor-bound conformations of IL-2 than the free forms, having undergone a relatively small shift towards the helical core as a consequence of the stabilizing mutations (Fig. 2c).

We used molecular dynamics simulations of IL-2 and D10 to further interrogate the mechanism responsible for higher binding affinity to IL-2R β by the IL-2 superkine (Fig. 2d, e). We constructed an atomically detailed Markov state model (MSM) to probe the relative conformational flexibility of IL-2 versus D10 directly. Analysis of the MSM clearly demonstrated that D10 was more stable than IL-2, and that IL-2 visited nearly twice as many clusters as D10. For example, the most populated state of D10 had an equilibrium probability of approximately 0.20, compared to approximately 0.05 for IL-2, demonstrating that the equilibrium population of D10 was far more localized than IL-2. Helix B, the B-C loop and helix C appeared rigidified in D10 compared to IL-2 as evidenced by reduced r.m.s.d. from the starting conformations (Fig. 2d and Supplementary Movies 1, 2). F92 seemed to act as a molecular wedge between helix C and helix A, stabilizing the more C-terminal end of the helix (Fig. 2a). We also simulated both D10 and IL-2 starting in a receptor-bound-like structure and monitored the divergence in r.m.s.d. of the B-C loop and helix C from the actual receptor-bound structure. IL-2 (Fig. 2e, left, and Supplementary Fig. 8c) quickly 'wandered' from the receptor conformation and experienced drastic fluctuations compared to D10 (Fig. 2e, right, Supplementary Fig. 8c and Supplementary Movies 1 and 2). Based on these observations, we propose a mechanism whereby the reduced flexibility of helix C in the IL-2 superkine, as a result of its improved core packing with helix B, results in a superior receptor-binding pose that increases its affinity for IL-2R β , and consequently mimics a functional role of CD25. The structural and molecular dynamics results indicate that the evolved mutations in the IL-2 superkine cause a conformational stabilization of the cytokine that reduces the energetic penalties for binding to IL-2R β .

We asked if the IL-2 superkines demonstrated signalling potencies on cells in accordance with their IL-2R β -binding affinities, and whether their activities depended on cell surface expression of CD25. We determined the dose-response relationships of wild-type IL-2 versus the IL-2 superkines 6-6, D10 and H9 on both CD25⁻ and CD25⁺ human YT-1 NK cells by assaying STAT5 phosphorylation with flow cytometry (Fig. 3a–d and Supplementary Fig. 9). On CD25⁻ YT-1 cells, the half-maximum effective concentration (EC₅₀) of H9 and D10 were decreased over tenfold (EC₅₀ = 2.5 and 1.8 ng ml⁻¹, respectively) compared to IL-2 (EC₅₀ = 39 ng ml⁻¹), with the 6-6 mutein yielding an EC₅₀ intermediate between IL-2 and H9/D10 (EC₅₀ = 15 ng ml⁻¹), consistent with the improved affinity of the IL-2 superkines for IL-2R β (Fig. 3a). On CD25⁺ YT-1 cells, the EC₅₀ of IL-2 decreased over 50-fold relative to CD25⁻ YT-1 cells, from 39 to 0.66 ng ml⁻¹ (Fig. 3b). In contrast, the EC₅₀ of H9 and D10 improved only modestly in the presence of CD25 (EC₅₀ of 0.47 and 0.52 compared to 2.5 and 1.8 ng ml⁻¹, respectively) (Fig. 3b).

We sought to further probe the CD25-independence of the IL-2 superkines by taking advantage of a previously characterized mutation

in IL-2, Phe 42 to Ala (F42A), which showed reduced binding to CD25 by approximately 220-fold for H9 (K_d 6.6 nM versus 1.4 μ M) and approximately 120-fold for IL-2 (K_d 6.6 nM versus 0.8 μ M) (Supplementary Fig. 10)^{12,13}. The F42A mutation is an alternative diagnostic probe of the relative CD25 dependence of IL-2 and the IL-2 superkine. The F42A mutation right-shifted the dose-response curve of wild-type IL-2 on CD25⁺ cells by about 1 log, but had no effect on CD25⁻ cells (Fig. 3c). In contrast, H9 was far less sensitive to the F42A mutation, with the dose-response curves of H9 versus H9 F42A being very similar on both CD25⁺ and CD25⁻ cells (Fig. 3d).

We assessed the activity of several IL-2 superkines on T cells that were either deficient in, or expressed CD25. For the former experiment, CD4⁺ T cells were isolated from CD25-knockout mice, followed by stimulation by either wild-type IL-2 or six IL-2 superkines and assaying for STAT5 phosphorylation at a range of cytokine concentrations (Fig. 3e and Supplementary Fig. 11). CD25^{-/-} CD4⁺ T cells responded poorly to exogenous wild-type IL-2 stimulation, but the IL-2 superkines induced STAT5 phosphorylation in these cells proportional to their affinity for IL-2R β .

The principle functional effect of IL-2 is to promote T cell proliferation, particularly for naive T cells. Human naive CD4⁺ T cells were isolated and left either unstimulated or stimulated with plate-bound anti-CD3 antibody with or without the different IL-2 variants (Fig. 3f and Supplementary Fig. 12). Increased proliferation effects on naive human T cells correlated with increased affinity for IL-2R β and STAT5 phosphorylation shown earlier, as the rank order of potency was D10 = H9 > 6-6 > wild-type IL-2 (see Supplementary Fig. 12 for the complete titration).

We next tested the IL-2 variants for their ability to induce STAT5 phosphorylation on experienced human CD4⁺ T cells (Supplementary Fig. 13), which highly express the trimeric IL-2R $\alpha\beta\gamma$ complex. Human CD4⁺ T cells were activated *in vitro* by T cell receptor (TCR) stimulation and rested to generate 'experienced' human CD4⁺ CD25⁺ T cells. As for the CD25⁺ YT-1 cells, we observed a much smaller difference between IL-2 and the IL-2 superkines.

We assessed the potency of the IL-2 superkine H9 on expansion of CD25^{low} versus CD25^{high} T cells, in comparison to wild-type IL-2 and IL-2-anti-IL-2 monoclonal antibody (mAb) complexes, which have been shown to exert reduced pulmonary oedema yet very potent antitumour responses *in vivo*^{14–16}. On antigen-experienced (memory-phenotype, MP) CD8⁺ T cells, expressing only low levels of CD25 but high levels of IL-2R $\beta\gamma$, H9 induced more than three times the rate of proliferation and expansion as wild-type IL-2 (Fig. 4a and Supplementary Fig. 14a). However, on CD4⁺ CD25^{high} T regulatory (T_{reg}) cells, we found that the CD25-competent wild-type IL-2 and H9 achieved comparable maximal expansion, demonstrating again that expression of CD25 mitigates the difference between the IL-2 superkine and wild-type IL-2 (Fig. 4a and Supplementary Fig. 14b). Thus, the H9 has the desired property that it shows enhanced stimulation towards CD8⁺ T cells, but not towards T_{reg} cells, compared to wild-type IL-2.

As previously reported, administration of high-dose wild-type IL-2 for 5 days induced substantial pulmonary oedema, which is known to be CD25-dependent¹⁵ (Fig. 4b). Although significantly more stimulatory for cytotoxic CD8⁺ T cells (Fig. 4a), the H9 IL-2 superkine caused substantially less pulmonary oedema (Fig. 4b).

Given the more favourable properties of H9 in comparison to IL-2, we assessed its ability to stimulate effector functions of cytotoxic T cells in four different tumour models *in vivo*, where high-dose IL-2 administration has been previously shown to result in tumour regression^{15,17}. To this end, C57BL/6 mice were injected subcutaneously with B16F10 melanoma cells, followed by administration of either high-dose IL-2, IL-2-anti-IL-2 mAb complexes, or the H9 IL-2 superkine, once tumour nodules became visible and palpable. PBS-treated control mice rapidly developed large subcutaneous tumours reaching a volume of about 1,500 mm³ on day 18 (Fig. 4c). As previously shown, high-dose IL-2 treatment was able to delay tumour growth by as much as 39% on

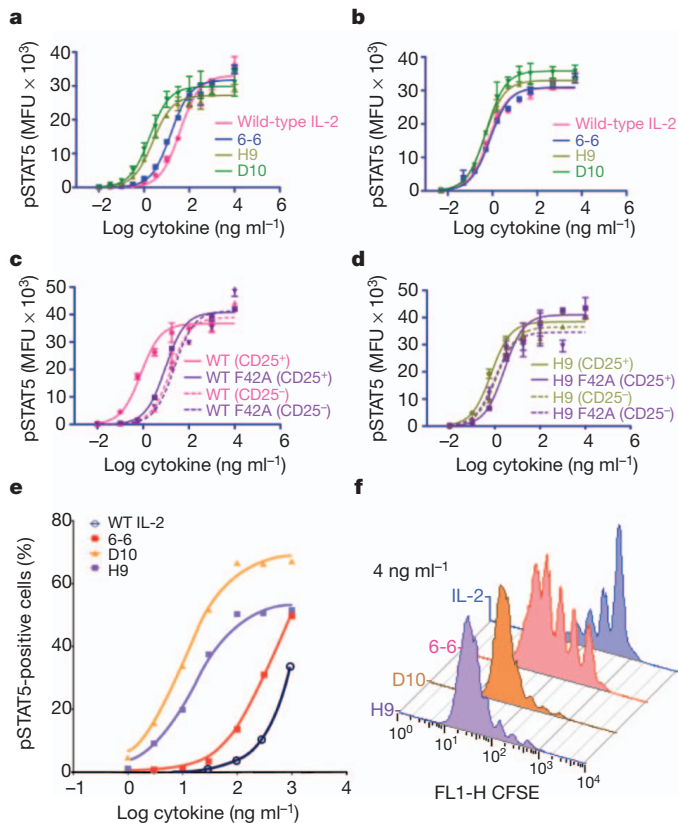


Figure 3 | Functional properties of the IL-2 superkine on human NK cells *in vitro*. **a, b**, Dose-response curves of STAT5 phosphorylation (pSTAT5) on CD25⁻ (**a**) and CD25⁺ (**b**) YT-1 cells with wild-type IL-2 and three IL-2 superkines. MFU, mean fluorescence units. **c**, Dose-response curves of STAT5 phosphorylation on CD25⁺ (solid curves) and CD25⁻ (dashed curves) YT-1 cells with wild-type IL-2 (pink curves) and IL-2-F42A mutation (purple curves). **d**, Dose-response curves of STAT5 phosphorylation on CD25⁺ (solid curves) and CD25⁻ (dashed curves) YT-1 cells with H9 (green curves) and H9-F42A mutation (purple curves). **e**, The IL-2 superkines have superior potency over IL-2 on T cells derived from CD25^{-/-} mice as demonstrated by dose-response curves for STAT5 phosphorylation on T cells demonstrating that potency correlates with IL-2R β affinity (see also Supplementary Fig. 10). **f**, Proliferation of human naive CD4⁺ T cells (CD25^{low}) reveals similar potency profiles as seen with CD25^{-/-} T cells. Proliferation was measured by carboxyfluorescein succinimidyl ester (CFSE) dilution on day 5 (see also Supplementary Fig. 10). Error bars in **a–d** represent s.e.m. of mean fluorescence units for each sample at the indicated cytokine concentration.

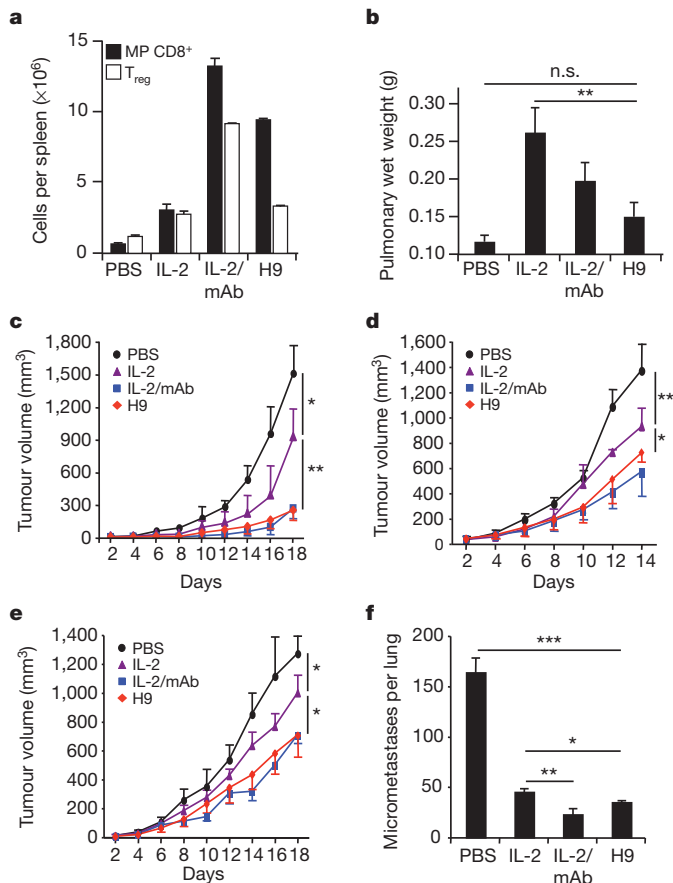


Figure 4 | Functional and antitumour activities of the IL-2 superkine *in vivo*. **a**, Total cell counts of host CD3⁺ CD8⁺ CD44^{high} memory-phenotype T cells (MP CD8⁺, closed bars), and host CD3⁺ CD4⁺ CD25^{high} T cells (T_{reg}, open bars) was determined in the spleens of mice receiving either PBS, 20 μ g IL-2, 1.5 μ g IL-2-anti-IL-2 mAb complexes (IL-2/mAb), or 20 μ g H9 (see also Supplementary Fig. 14). **b**, Pulmonary oedema (pulmonary wet weight) served to assess adverse toxic effects following IL-2 treatment, and was determined by weighing lungs before and after drying. **c–f**, C57BL/6 mice ($n = 3–4$ mice per group) were injected either subcutaneously with 10⁶ B16F10 melanoma cells (**c**), 2.5 $\times 10^6$ murine colon carcinoma 38 (**d**), 10⁶ Lewis lung carcinoma (**e**), or mice received 3 $\times 10^5$ B16F10 melanoma cells intravenously (**f**), followed by daily injections of either PBS, 20 μ g IL-2, 1.5 μ g IL-2/mAb complexes, or 20 μ g H9 for 5 days once subcutaneous tumour nodules became visible and palpable or from day three on for intravenously-injected tumours (see also Supplementary Fig. 15). Shown is mean tumour volume in mm³ (\pm s.d.) versus time upon tumour inoculation. Error bars represent s.e.m., P values refer to comparisons of wild type with the other treatment modalities. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

day 18 ($P < 0.05$), whereas IL-2-anti-IL-2 mAb complexes exerted very effective tumour control, reducing tumour growth by more than 80% on day 18 ($P < 0.005$) (Fig. 4c). Significantly, similar to IL-2-anti-IL-2 mAb complexes, mice receiving high-dose H9 showed a dramatic decrease of tumour load on day 18, which was reduced by more than 80% compared to PBS ($P < 0.005$) and by more than 70% compared to wild-type IL-2 ($P < 0.001$) (Fig. 4c). Similar results were obtained using three other tumour models, including murine colon carcinoma and Lewis lung carcinoma injected subcutaneously (Fig. 4d, e) and B16F10 cells administered intravenously (Fig. 4f and Supplementary Fig. 15). Collectively, these data show that the H9 IL-2 superkine is very effective against different tumours, albeit inducing reduced pulmonary oedema.

The practical implications are that this conformational nuance in IL-2 can be exploited for therapy. The IL-2 superkine robustly activates cytotoxic CD8⁺ T cells and NK cells for potent antitumour immune

responses, yet it elicits minimal toxicity, suggesting that the IL-2 superkine could warrant reconsideration for clinical applications of IL-2.

METHODS SUMMARY

Yeast display and selection of IL-2. Error-prone and site-directed libraries of IL-2 were displayed on yeast as previously described¹⁸ and stained with biotinylated IL-2R β at successively decreasing concentrations. Staining was detected with streptavidin-phycoerythrin and yeast were separated using paramagnetic anti-phycoerythrin microbeads (Miltenyi; MACS). Enrichment of positively-staining yeast was monitored by flow-cytometry.

Protein expression, purification and structural determination. Human IL-2 variants and the ectodomains of IL-2R β , IL-2R γ and CD25 were expressed in Hi5 cells and purified as previously described¹¹. Proteins were concentrated to 8–20 mg ml⁻¹ and crystallized by vapour diffusion in sitting drops. Diffraction studies were performed at the Stanford Synchrotron Radiation Laboratory and the Advanced Light Source. Crystal structures were solved by molecular replacement with PHASER¹⁹ and refined using PHENIX²⁰ and COOT²¹.

Mice. C57BL/6 and Thy1.1-congenic mice on a C57BL/6 background were maintained under specific pathogen-free conditions and used at 3–6 months of age. Experiments were performed in accordance with the Swiss Federal Veterinary Office guidelines and approved by the Cantonal Veterinary Office.

***In vivo* T-cell proliferation.** Carboxyfluorescein succinimidyl ester (CFSE)-labelled CD44^{high} CD8⁺ T cells (2×10^6 to 3×10^6) from Thy1.1-congenic mice were injected intravenously to Thy1.2-congenic animals. Mice received daily intraperitoneal (i.p.) injections of PBS, 20 μ g IL-2, 1.5 μ g IL-2-anti-IL-2 mAb complexes, or 20 μ g H9 for 5 days. On the sixth day, spleens were removed and analysed by flow cytometry.

Toxicity. Pulmonary oedema was determined by measurement of pulmonary wet weight on the sixth day after five daily i.p. injections of PBS, 20 μ g IL-2, 1.5 μ g IL-2-anti-IL-2 mAb complexes, or 20 μ g H9 as previously described¹⁵.

Tumour models. B16F10 melanoma cells, Lewis lung carcinoma or murine colon carcinoma 38 cells were injected into mice (3–4 mice per group), as previously reported^{15,17}. Treatment consisted of five daily i.p. injections of PBS, 20 μ g IL-2, 1.5 μ g IL-2-anti-IL-2 mAb complexes, or 20 μ g H9.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 October 2011; accepted 20 February 2012.

Published online 25 March 2012.

- Rochman, Y., Spolski, R. & Leonard, W. J. New insights into the regulation of T cells by γ_c family cytokines. *Nature Rev. Immunol.* **9**, 480–490 (2009).
- Smith, K. A. Interleukin-2: inception, impact, and implications. *Science* **240**, 1169–1176 (1988).
- Waldmann, T. A. The biology of interleukin-2 and interleukin-15: implications for cancer therapy and vaccine design. *Nature Rev. Immunol.* **6**, 595–601 (2006).
- Cosman, D. *et al.* Cloning, sequence and expression of human interleukin-2 receptor. *Nature* **312**, 768–771 (1984).
- Leonard, W. J. *et al.* Molecular cloning and expression of cDNAs for the human interleukin-2 receptor. *Nature* **311**, 626–631 (1984).
- Nikaido, T. *et al.* Molecular cloning of cDNA encoding human interleukin-2 receptor. *Nature* **311**, 631–635 (1984).
- Hatakeyama, M. *et al.* Interleukin-2 receptor beta chain gene: generation of three receptor forms by cloned human alpha and beta chain cDNAs. *Science* **244**, 551–556 (1989).
- Takeshita, T. *et al.* Cloning of the gamma chain of the human IL-2 receptor. *Science* **257**, 379–382 (1992).
- Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnol.* **15**, 553–557 (1997).
- Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nature Protocols* **1**, 755–768 (2006).
- Wang, X., Rickert, M. & Garcia, K. C. Structure of the quaternary complex of interleukin-2 with its α , β , and γ_c receptors. *Science* **310**, 1159–1163 (2005).
- Mott, H. R. *et al.* The solution structure of the F42A mutant of human interleukin 2. *J. Mol. Biol.* **247**, 979–994 (1995).
- Thanos, C. D., DeLano, W. L. & Wells, J. A. Hot-spot mimicry of a cytokine receptor by a small molecule. *Proc. Natl Acad. Sci. USA* **103**, 15422–15427 (2006).
- Boyman, O., Kovar, M., Rubinstein, M. P., Surh, C. D. & Sprent, J. Selective stimulation of T cell subsets with antibody-cytokine immune complexes. *Science* **311**, 1924–1927 (2006).
- Krieg, C., Létourneau, S., Pantaleo, G. & Boyman, O. Improved IL-2 immunotherapy by selective stimulation of IL-2 receptors on lymphocytes and endothelial cells. *Proc. Natl Acad. Sci. USA* **107**, 11906–11911 (2010).
- Létourneau, S. *et al.* IL-2/anti-IL-2 antibody complexes show strong biological activity by avoiding interaction with IL-2 receptor α subunit CD25. *Proc. Natl Acad. Sci. USA* **107**, 2171–2176 (2010).

17. Rosenberg, S. A., Mule, J. J., Spiess, P. J., Reichert, C. M. & Schwarz, S. L. Regression of established pulmonary metastases and subcutaneous tumor mediated by the systemic administration of high-dose recombinant interleukin 2. *J. Exp. Med.* **161**, 1169–1188 (1985).
18. Rao, B. M., Driver, I., Lauffenburger, D. A. & Wittrup, K. D. Interleukin 2 (IL-2) variants engineered for increased IL-2 receptor α -subunit affinity exhibit increased potency arising from a cell surface ligand reservoir effect. *Mol. Pharmacol.* **66**, 864–869 (2004).
19. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
20. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
21. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors gratefully acknowledge W. Leonard, R. Levy and R. Schwendener for reagents and discussion. This work was supported by NIH-R01AI51321 (to K.C.G.), PP00P3-128421 from the Swiss National Science Foundation and KFS-02672-08-2010 from the Swiss Cancer League (both to O.B.), NIH R01-GM062868 (to V.S.P.), MRI-R2 (this award is funded under the American Recovery and Reinvestment Act of 2009 (Public Law 111-5)) (to V.S.P.),

NIH-AR050942 (to J.T.L.), NIH U01 DK078123 (to C.G.F.), and NIH U19 AI 082719 (to C.G.F.). A.M.R. was supported by the Stanford Medical Scientist Training Program (NIH-GM07365). K.C.G. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions A.M.L. performed *in vitro* evolution and contributed to preparation of the manuscript. D.L.B. produced recombinant proteins, determined crystal structures, and carried out surface plasmon resonance analysis. A.M.R. carried out cellular and signalling assays, biophysical measurements and contributed to preparation of the manuscript. C.K. carried out *in vivo* experiments, analysed data and contributed to preparation of the manuscript; M.E.R. carried out *in vivo* experiments in mice. I.M. analysed cell-signalling data. G.R.B., P.N. and V.S.P. carried out and analysed molecular dynamics simulations. J.T.L., L.S. and C.G.F. performed and analysed T-cell signalling experiments. O.B. designed and supervised *in vivo* experiments, analysed data and contributed to preparation of the manuscript. K.C.G. conceived of the project, analysed data, supervised execution of the project, and prepared the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession codes 3QAZ and 3QB1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.C.G. (kcgarcia@stanford.edu) or O.B. (onur.boyman@uzh.ch).

METHODS

Yeast display of wild-type IL-2. Human IL-2 DNA was cloned into the vector pCT302 and displayed on the *Saccharomyces cerevisiae* strain EBY100 as previously described¹⁸. Individual colonies of IL-2 yeast were grown overnight at 30 °C in SDCAA liquid media and induced in SGCAA liquid media for 2 days at 20 °C. The yeast were stained with tetramerized biotinylated IL-2R β (b-IL-2R β), biotinylated IL-2R γ (b-IL-2R γ), or b-IL-2R β in the presence of b-IL-2R γ . IL-2R β tetramers were formed by incubating 2 μ M b-IL-2R β with 470 nM SAV-PE (streptavidin–phycoerythrin conjugate, Invitrogen) in phosphate buffered saline supplemented with 0.5% BSA and 2 mM EDTA (PBE) for 15 min on ice. Analysis was performed on an Accuri C6 flow cytometer.

Error-prone PCR IL-2 library construction and selection. Human IL-2 DNA was subjected to error-prone PCR using the GeneMorphII kit (Stratagene). The two user-determined variables in the kit were the starting concentration of DNA template and the number of cycles. We used 100 ng template and 30 cycles in an effort to maximize the number of errors. The primers used for error-prone PCR were 5'-GCACCTACTTCAAGTTCTAC-3' for the forward primer and 5'-GC CACCAGAGGATCC-3' for the reverse primer. The PCR product was further amplified using primers containing sequence homology to pCT302 for recombination inside the yeast: forward primer 5'-AGTGGTGGTGGTGGTCTCTGG TGGTGGTGGTCTGGTGGTGGTGGTCTGCTAGCGCACCTACTTCAAG TTCTAC-3' and reverse primer 5'-ACACTGTTGTATCAGATCTCGAGCAA GTCTTCTTCGGAGATAAGCTTTTGTTCGCCACAGAGGATCC-3'. Insert DNA was combined with linearized vector backbone and EBY100 yeast and electroporated as previously described¹⁰. The electroporations yielded a library of 1×10^8 transformants. Selections were performed using magnetic activated cell sorting (MACS, Miltenyi). First-round selection was performed with 1×10^9 yeast, approximately tenfold coverage of the number of transformants. Yeast were stained with 10 ml of 500 nM IL-2R β SAV-PE tetramers in PBE for 2 h with slow rotation at 4 °C. Cells were pelleted at 5,000g for 5 min, buffer aspirated, and washed with PBE. The pellet was resuspended in 9.6 ml PBE and 400 μ l Miltenyi anti-PE microbeads, incubated for 20 min with slow rotation at 4 °C, pelleted, and washed with 14 ml PBE. The yeast were then resuspended in 5 ml PBE and magnetically separated by a Miltenyi LS column, following the manufacturer's protocols. Subsequent rounds of selection used 1×10^8 yeast cells, successively lower concentrations of monomeric IL-2R β for increased selection stringency, and 100 μ l of Miltenyi microbeads to capture labelled yeast-displayed IL-2 variants.

Site-directed hydrophobic core IL-2 library construction and selection. The site-directed library was created by assembly PCR of 13 primers, two of which contained the following degenerate codons: Q74 = MRS, L80 = NTC, R81 = NNK, L85 = NTC, I86 = NTC, I89 = NTC, I92 = NTC, V93 = NTC. The primers used for assembly PCR were: 5'-GCACCTACTTCAAGTTCTACAA AGAAACACAGCTACAAGTGGAGCA-3', 5'-CAAAATCATCTGTAAATC CAGAAGTAAATGCTCCAGTTGTAGCTGTG-3', 5'-GGATTTACAGATGA TTTTGAATGGAATTAATAATTACAAGAATCCCA-3', 5'-AACTTAGCTGT GAGCATCCTGGTGAGTTTGGGATCTTGTAAATTATT-3', 5'-GGATGCTC ACAGCTAAGTTTACATGCCCAAGAGGCCACAGAAGT-3', 5'-GTTCT TCTCTAGACACTGAGATGTTTACAGTCTGTGGCCTTCTGTG-3', 5'-CA GTGTCTAGAAGAAGAACTCAAACCTCTGGAGGAAGTGCTAAATTTA-3', 5'-GTGAAAGTTTGTCTSYKAGCTAAATTTAGCACTTCTCTCC-3', 5'-AG CAAAACTTTTCACNTCNKCCAGGAGCANTCNCAGCAATNTCAACG TANTACTCTGGAAGTAAAGGGATG-3', 5'-CATCAGCATATTCACACA TGAATGTTGTTTACAGATCCCTTTAGTTCAG-3', 5'-ATGTGTGAATATG CTGATGAGACAGCAACCATTGTAGAATTTCTGAACA-3', 5'-AGATGAT GCTTTGACAAAAGGTAATCCATCTGTTCAGAAATTCTACAAT-3', 5'-TT TTGTCAAAGCATCATCTCAACACTAACTGGATCCTCTGGTGGC-3'. The assembly PCR reaction was performed using *Pfu* DNA polymerase (Stratagene). The product DNA was further PCR-amplified using the same primers as the error-prone library. Electroporation of insert DNA and linearized vector into EBY-100 yeast yielded a library of 1.4×10^8 transformants. Selection of the site-directed library was performed as with the error-prone library, except tenfold lower concentrations of monomeric IL-2R β were used.

Protein expression and purification. Human IL-2 variants (amino acids 1–133), the IL-2R β ectodomain (amino acids 1–214), IL-2R γ (amino acids 34–232) and CD25 (amino acids 1–217) were expressed and purified from Hi5 cells as previously described¹¹. For biotinylated receptor expression, IL-2R β and IL-2R γ with a C-terminal biotin acceptor peptide (BAP)-LNDIFEAKIEWHE were co-expressed with BirA ligase with excess biotin (100 μ M). CD25 was biotinylated at its free cysteine by incubation with biotin–maleimide (Sigma). For crystallization, receptor constructs were co-expressed with N-linked glycosylation sites IL-2R β residues Asn 3, 17, 45, and IL-2R γ residues Asn 53 mutated to Gln. IL-2 proteins used for crystallization were expressed with N-linked glycosylation

inhibitor tunicamycin (0.2 μ g ml⁻¹). Proteins were treated overnight with carboxypeptidase-A followed by size-exclusion chromatography.

Surface plasmon resonance. SPR experiments were conducted on a Biacore T100 instrument at 25 °C. All data was analysed using the Biacore T100 evaluation software version 2.0 with a 1:1 Langmuir binding model. Experiments used a Biacore SA sensor chip (GE Healthcare). Biotinylated receptors were captured at a low density ($R_{\max} \approx 30$ response units) and kinetic runs were conducted at 40 μ l min⁻¹ to eliminate mass transport and rebinding artefacts. An unrelated biotinylated protein was immobilized as a reference surface. All measurements were made using threefold serial dilutions of IL-2 variants (GE Healthcare, 0.01% BSA). IL-2R β was regenerated using 10 mM sodium acetate (pH 5.5) and 1 M MgCl₂. CD25 was regenerated using 1 M NaCl. Kinetic data was determined using 120 s to 190 s of IL-2 variant association and 20 s to 600 s dissociation.

Isothermal titration calorimetry. Calorimetry experiments with the H9 and D10 IL-2 superkinetics with IL-2R β were conducted as previously described²². Briefly, titrations were performed on a VP-ITC calorimeter (MicroCal) at 15 °C. Prior to titration, protein and reference water was degassed for 10 min. All samples were extensively dialysed in 10 mM HEPES pH 7.4, 150 mM NaCl (HBS) before titration to minimize heat of dilution effects. Data were processed and analysed using the MicroCal Origin 5.0 software.

Crystallization and data collection. IL-2 D10 crystals were grown in sitting drops at 22 °C from 50 mM HEPES (pH 7.2), 200 mM NaCl and 30% PEG-4000. A 3.1 Å data set was collected under cryo-cooled conditions (20% glycerol) at beamline 11-1 at the Stanford Synchrotron Radiation Laboratory. IL-2 D10 ternary crystals were grown from 100 mM Bis-Tris (pH 5.5), 200 mM NH₄SO₄ and 25% PEG-3350. A cryo-cooled 3.8 Å data set was collected at beamline 8-2 at the Advanced Light Source. Diffraction data were processed using HKL2000. Data processing statistics can be found in Supplementary Table 1.

Structure determination and refinement. The IL-2 D10 and IL-2 D10 ternary crystal structures were solved by molecular replacement with the program PHASER¹⁹ using the coordinates of IL-2 (PDB ID 1M47) and the quaternary complex (PDB ID 2B5I), respectively, and refined with PHENIX²⁰ and COOT²¹ (Supplementary Table 1). Bulk solvent flattening was used for solvent correction in both structures. For both the IL-2 D10 free structure and the ternary complex non-crystallographic symmetry (NCS) restraints (not constraints) were used for initial stages of the refinement. Coordinate refinement strategies included rigid body, restrained individual, group atomic displacement parameters (ADP) and torsion-simulated annealing. The final rounds of refinement removed all NCS restraints for minimization and a round of individual ADP refinement.

Ramachandran analysis was performed with MolProbity²³. Buried surface area values were calculated using the Protein Interfaces, Surfaces, and Assemblies (PISA) software²⁴. IL-2 D10 consisted of eight chains with chain A displayed in the paper. The IL-2 D10 ternary complex contained 36 chains comprising 12 ternary complexes. The paper figures are from chains A, B and C. All structural figures and overlays were prepared using PyMOL²⁵.

Tissue culture and magnetic purification of CD25⁺ YT-1 cells. YT-1 and CD25⁺ YT-1 cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, minimum non-essential amino acids, sodium pyruvate, 25 mM HEPES, and penicillin-streptomycin (Gibco). CD25⁺ YT-1 cells were purified as follows: 1×10^7 cells were washed with FACS buffer (phosphate buffered saline + 2% bovine serum albumin) and stained with PE-conjugated anti-human CD25 (1:20; Biotend) in 1 ml FACS buffer for 20 min at 4 °C. The stained cells were labelled with 200 μ l paramagnetic microbeads coupled to anti-PE IgG and separated with an LS MACS separation column according to the manufacturer's instructions (Miltenyi Biotec). Eluted cells were re-suspended in complete RPMI medium at a concentration of 1×10^5 cells per ml and expanded for subsequent experiments. Enrichment of cells was monitored via flow cytometry with the FL-2 channel using an Accuri C6 flow cytometer.

YT-1 dose-response experiments and phospho-flow cytometric analysis. 2×10^5 CD25⁺ or CD25⁻ YT-1 cells were washed with FACS buffer and re-suspended in 200 μ l FACS buffer with the indicated concentration of IL-2 variant per well in a 96-well plate. Cells were stimulated for 20 min at room temperature and then fixed by addition of formaldehyde to 1.5% and incubated for 10 min. Cells were permeabilized with 100% ice-cold methanol for 20 min on ice, followed by incubation at -80 °C overnight. Fixed, permeabilized cells were washed with excess FACS buffer and incubated with 50 μ l Alexa647-conjugated anti-STAT5 pY694 (BD Biosciences) diluted 1:20 in FACS buffer for 20 min. Cells were washed twice in FACS buffer and mean cell fluorescence determined using the FL-4 channel of an Accuri C6 flow cytometer. Dose-response curves and EC₅₀ values were calculated in GraphPad Prism after subtracting the mean cell fluorescence of unstimulated cells.

For 'on-yeast' stimulation experiments, the same protocol was used with the following modifications. Induced yeast were washed twice in FACS buffer and

mixed with the 2×10^5 YT cells at the indicated ratios for 20 min in FACS buffer at room temperature. Cells were then fixed, permeabilized and stained as described above.

T cell isolation and proliferation for phospho-flow cytometric analysis. Human and mouse CD4 T cells were prepared from peripheral blood mononuclear cells (PBMC, Stanford Blood Bank) and spleens and lymph nodes of BALB/C mice, respectively, using antibody-coated CD4 T-cell isolation magnetic beads (Stem Cell Technologies and Miltenyi Biotec). For naive cell stimulation assays, cells were used immediately. For generation of *in vitro* 'experienced' T cells, wells were pre-coated with secondary antibody (Vector Labs) in bicarbonate buffer, pH 9.6 before coating plates with anti-CD3 (OKT3 for human, 2C11 for mouse, eBiosciences) at 100 ng ml⁻¹. T cells were seeded at 0.1×10^6 cells per well with soluble anti-CD28 (CD28.2 for human, 37.51 for mouse, eBiosciences). Cells were cultured for 3 days with full T-cell receptor stimulation, followed by 2 days rest in conditioned media and 2 days rest in fresh culture media. Prior to use, live cells were collected by Lympholyte-M (Cederlane) centrifugation and counted.

In vivo studies. C57BL/6 and Thy1.1-congenic mice on a C57BL/6 background (both from Charles River) were maintained under specific pathogen-free conditions and used at 3–6 months of age. Experiments were performed in accordance with the Swiss Federal Veterinary Office guidelines and approved by the Cantonal Veterinary Office.

Cell suspensions of spleen were prepared according to standard protocols¹⁴ and stained for analysis by flow cytometry using phosphate-buffered saline (PBS) containing 4% fetal calf serum and 2.5 mM EDTA. Fluorochrome-conjugated monoclonal antibodies (mAbs) (from BD Biosciences unless otherwise stated) were used against: CD3 (145-2C11, eBioscience), CD4 (RM4-5, Caltag Laboratories), CD8a (53-6.7), CD25 (PC61), CD44 (IM7, eBioscience), NK1.1 (PK136), and Thy1.1 (HIS51, eBioscience). At least 100,000 viable cells were acquired on a BD FACSCanto II flow cytometer and analysed using FlowJo software (TriStar Inc.).

To prepare IL-2-anti-IL-2 mAb complexes, recombinant human IL-2 (rhIL-2) and anti-human IL-2 mAb were premixed at a 2:1 molar ratio using 15,000 international units of recombinant human IL-2, as previously described¹⁵. Recombinant human IL-2 and anti-human IL-2 mAb clone 5355 (MAB602) were obtained from R&D Systems.

T-cell subsets were obtained by negative T-cell enrichment (StemCell Technologies). Where indicated, purified cells were labelled with carboxyfluorescein diacetate succinimidyl ester (CFSE, Molecular Probes), as previously published¹⁴. 2×10^6 to 3×10^6 CD8⁺ T cells from Thy1.1-congenic wild-type mice enriched for CD44^{high} memory-phenotype cells were injected intravenously (i.v.) to Thy1.2-congenic wild-type mice. Starting on the day of adoptive cell transfer, age- and gender-matched mice received daily intraperitoneal (i.p.) injections of either PBS, 20 µg wild-type human IL-2, 20 µg H9, or 1.5 µg human IL-2-anti-IL-2 mAb complexes for 5 consecutive days. 6 days after adoptive cell transfer, spleens were removed and analysed by flow cytometry.

Pulmonary wet weight was determined according to previously established protocols¹⁵. In brief, wild-type mice received daily i.p. injections of either PBS, 20 µg wild-type human IL-2, 20 µg H9, or 1.5 µg IL-2-anti-IL-2 mAb complexes for 5 consecutive days as described above. On day 6, lungs were removed and weighed before and after drying overnight at 58 °C under vacuum. Pulmonary wet weight was calculated by subtracting initial pulmonary weight from lung weight after dehydration.

To generate subcutaneous tumours, as indicated either 10⁶ B16F10 melanoma (from ATCC), 10⁶ Lewis lung carcinoma (provided by R. Schwendener), or 2.5×10^6 murine colon carcinoma 38 (provided by R. Schwendener) cells were injected in 100 µl DMEM into the upper dermis of the back of mice (3–4 mice per group), as previously established¹⁵. Treatment consisted of five daily injections of either PBS, 20 µg IL-2, 1.5 µg IL-2-anti-IL-2 mAb complexes, or 20 µg H9, and was started 1 day after tumour nodules were clearly visible and palpable at a volume of approximately 50–55 mm³. For the generation of lung metastases, 3×10^5 B16F10 cells in 300 µl DMEM were injected into the tail vein, as previously shown¹⁵. Treatment was as above and was started on day 3 after tumour inoculation. On day 16 after injection, lungs were perfused, harvested and fixed in Fekete's solution (70% ethanol, 3.7% paraformaldehyde, 0.75 M glacial acetic acid), followed by dissection of lungs and counting of pulmonary micrometastases. Differences between groups were examined for statistical significance by using a one-way analysis of variance (ANOVA) with Bonferroni post-test correction.

Molecular dynamics simulations and MSM. We used MODELLER²⁶ with the default settings to create five starting conformations for simulations from each of

three IL-2 structures (PDB ID 3INK, 1M47 and 1Z92; refs 27–29), and D10 with both the wild-type IL-2 sequence and the D10 sequence. The five conformations for a particular sequence/structure differ only where residues are missing or mutated.

Molecular dynamics simulations were run with Gromacs 4.5.2 (ref. 30) using the AMBER03 force field³¹. Each structure was placed in a dodecahedral box of about 7.1 by 7.1 by 5 nm and solvated with approximately 7,650 TIP3P water molecules. Conformations were first minimized with a steepest descent algorithm using a tolerance of 1,000 kJ mol⁻¹ nm⁻¹ and a step size of 0.01 nm. A 1-nm cutoff was used for Coulombic and Van der Waals interactions and a grid-based neighbour list. Conformations were then equilibrated at 300 K and 1 bar by holding protein atoms fixed and allowing the surrounding water to relax for 500 ps with a 2 fs time-step. All bonds were constrained with the LINCS algorithm³². Centre of mass motion was removed at every step and a grid-based neighbour list with a cutoff of 1.5 nm was updated every 10 steps. For electrostatics, we used fourth order PME³³ with a cutoff of 1.5 nm for Coulombic interactions, a Fourier spacing of 0.08 nm, and a tolerance of 1×10^{-5} . A hard cutoff of 1.2 nm was used for Van der Waals interactions with a switch starting at 1 nm. The temperature was controlled with two Nose-Hoover thermostats³⁴ applied to the protein and solvent respectively with a time constant of 0.5 ps. The pressure was controlled with an isotropic Berenson barostat³⁵ applied to the entire system with a time constant of 0.5 ps and a compressibility of 4.5×10^{-5} bar⁻¹. Long-range corrections were applied to energy and pressure. Production simulations up to 40 ns duration used the same parameters as for equilibration, with the exception that the protein atoms were no longer held fixed.

We used MSMBuilder³⁶ to construct an MSM with a 4-ns lag time. Based on previous work on protein folding³⁷, we chose to create 70 clusters (microstates) using a k-centres algorithm and the r.m.s.d. between pairs of conformations. All C_α and C_β atoms were used for the r.m.s.d., thereby allowing different sequences to be used in the same clustering. Thermodynamic and kinetic properties were extracted from the MSM's eigenvalues and eigenvectors^{38,39}.

22. Rickert, M., Boulanger, M. J., Goriatcheva, N. & Garcia, K. C. Compensatory energetic mechanisms mediating the assembly of signaling complexes between interleukin-2 and its α , β , and γ_c receptors. *J. Mol. Biol.* **339**, 1115–1128 (2004).
23. Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
24. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
25. DeLano, W. L. The PyMOL Molecular Graphics System (DeLano Scientific, 2002).
26. Eswar, N. *et al.* Comparative Protein Structure Modeling With MODELLER. *Current Protocols in Bioinformatics*. Vol. 15 5.6.1–5.6.30 (Wiley, 2006).
27. Bazan, J. F. Unraveling the structure of IL-2. *Science* **257**, 410–413 (1992).
28. Arkin, M. R. *et al.* Binding of small molecules to an adaptive protein-protein interface. *Proc. Natl Acad. Sci. USA* **100**, 1603–1608 (2003).
29. Rickert, M., Wang, X., Boulanger, M. J., Goriatcheva, N. & Garcia, K. C. The structure of interleukin-2 complexed with its alpha receptor. *Science* **308**, 1477–1480 (2005).
30. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).
31. Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012 (2003).
32. Hess, B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).
33. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
34. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
35. Berendsen, H. J. C., Postma, P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
36. Bowman, G. R., Huang, X. & Pande, V. S. Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* **49**, 197–201 (2009).
37. Bowman, G. R., Beauchamp, K. A., Boxer, G. & Pande, V. S. Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.* **131**, 124101 (2009).
38. Bowman, G. R., Huang, X. & Pande, V. S. Network models for molecular kinetics and their initial applications to human health. *Cell Res.* **20**, 622–630 (2010).
39. Noé, F. & Fischer, S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol.* **18**, 154–162 (2008).

CORRIGENDUM

doi:10.1038/nature11026

Corrigendum: Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma

Jeremy Schwartzentruber, Andrey Korshunov, Xiao-Yang Liu, David T. W. Jones, Elke Pfaff, Karine Jacob, Dominik Sturm, Adam M. Fontebasso, Dong-Anh Khuong Quang, Martje Tönjes, Volker Hovestadt, Steffen Albrecht, Marcel Kool, Andre Nantel, Carolin Konermann, Anders Lindroth, Natalie Jäger, Tobias Rausch, Marina Ryzhova, Jan O. Korb, Thomas Hielscher, Peter Hauser, Miklos Garami, Almos Klekner, Laszlo Bogner, Martin Ebinger, Martin U. Schuhmann, Wolfram Scheurlen, Arnulf Pekrun, Michael C. Frühwald, Wolfgang Roggendorf, Christoph Kramm, Matthias Dürken, Jeffrey Atkinson, Pierre Lepage, Alexandre Montpetit, Magdalena Zakrzewska, Krzysztof Zakrzewski, Pawel P. Liberski, Zhifeng Dong, Peter Siegel, Andreas E. Kulozik, Marc Zapatka, Abhijit Guha, David Malkin, Jörg Felsberg, Guido Reifenberger, Andreas von Deimling, Koichi Ichimura, V. Peter Collins, Hendrik Witt, Till Milde, Olaf Witt, Cindy Zhang, Pedro Castelo-Branco, Peter Lichter, Damien Faury, Uri Tabori, Christoph Plass, Jacek Majewski, Stefan M. Pfister & Nada Jabado

Nature **482**, 226–231 (2012).

Sequence reads for paediatric glioblastoma multiforme (GBM) samples have been deposited in the European Genome Archive under the accession number EGAS00001000226. This has been corrected in the HTML and PDF versions online.

CORRIGENDUM

doi:10.1038/nature11027

Corrigendum: Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma

Ignacio Varela, Patrick Tarpey, Keiran Raine, Dachuan Huang, Choon Kiat Ong, Philip Stephens, Helen Davies, David Jones, Meng-Lay Lin, Jon Teague, Graham Bignell, Adam Butler, Juok Cho, Gillian L. Dalglish, Danushka Galappaththige, Chris Greenman, Claire Hardy, Mingming Jia, Calli Latimer, KingWai Lau, John Marshall, Stuart McLaren, Andrew Menzies, Laura Mudie, Lucy Stebbings, David A. Largaespada, L. F. A. Wessels, Stephane Richard, Richard J. Kahnoski, John Anema, David A. Tuveson, Pedro A. Perez-Mancera, Ville Mustonen, Andrej Fischer, David J. Adams, Alistair Rust, Waraporn Chan-on, Chutima Subimerb, Karl Dykema, Kyle Furge, Peter J. Campbell, Bin Tean Teh, Michael R. Stratton & P. Andrew Futreal

Nature **469**, 539–542 (2011).

In this Letter, the accession number for the gene expression data in the Gene Expression Omnibus was wrongly published as GEO22316. The correct accession number is GSE22316. We thank P. Boutros for pointing this out. This has been corrected in the HTML and PDF versions online.

GEOSCIENCE

Fossil raindrops and ancient air

An analysis of fossil imprints of ancient raindrops suggests that the density of the atmosphere 2.7 billion years ago was much the same as that today. This result casts fresh light on a long-standing palaeoclimate paradox.

WILLIAM S. CASSATA
& PAUL R. RENNE

When a main-sequence star such as our Sun ages, its inner core becomes denser and the temperature at which its hydrogen is fused to helium increases. As a result, the Sun is currently more luminous, and delivers more energy to Earth's surface, than in the past — its energy output around 2 billion years ago is inferred¹ to have been less than 85% of that today. Such a faint Sun should not have been able to warm Earth's surface above the freezing point of water², yet the geological record indicates that liquid water abounded at that time. The apparent contradiction between theory (sub-freezing Earth surface temperatures) and observation (liquid water) is known as the 'faint young Sun' paradox. In a paper published on *Nature's* website today, Som *et al.*³ address this paradox using seemingly unlikely evidence: fossilized imprints left by ancient raindrops.

There are ample indications of liquid water under the faint Sun. In South Africa, for example, exposed rocks more than 3 billion years old contain features associated with flowing or standing water, including sedimentary deposits that preserve ripple marks and mud cracks⁴, glassy 'pillow' lavas that were rapidly quenched⁵ and alga-like microfossils⁵. Indirect evidence, meanwhile, extends the record of terrestrial oceans to as early as 4.4 billion years ago⁶. So how can this be explained?

Solutions to the paradox fall into two general categories: those contending that Earth's atmosphere retained heat more efficiently in the past than it does now, for example because of increased concentrations of greenhouse gases, and those arguing that the albedo, or reflectance, of Earth was lower in the past, perhaps because there were fewer clouds and/or less ice. Most models used to explain the paradox are purely theoretical, and were designed to highlight the conditions necessary to solve it. Unfortunately, few observational constraints to support or refute these models have been identified, and those that have



Figure 1 | Solid evidence. Som *et al.*³ have analysed the fossilized imprints of raindrops, such as those shown here, to determine the atmospheric density 2.7 billion years ago. Rule, 5 cm.

been proposed^{7–10} tend to be controversial.

Som *et al.*³ have implemented a clever approach — first attempted¹¹ in 1851 by a pioneer of geology, Charles Lyell — to determine the density of the atmosphere early in Earth's history. This information is crucial for assessing whether greater concentrations of greenhouse gases (such as carbon dioxide¹²), or of other gases (such as nitrogen¹³) that amplify the effects of greenhouse gases, could explain the faint young Sun paradox.

Specifically, the authors inferred the velocity of falling ancient raindrops from the geometry of fossilized raindrop-impact marks preserved in a 2.7-billion-year-old sedimentary rock from South Africa (Fig. 1). The atmosphere exerts a drag on raindrops such that they typically fall at a terminal velocity that is inversely proportional to the density of the atmosphere. Any difference between the inferred velocity of ancient raindrops and that of those that fall today may therefore reflect a change in atmospheric density. Of course, the imprint generated by a raindrop falling at a given velocity depends

both on the size of the drop and on the nature of the substrate onto which it falls. To constrain these variables, the authors observed the size distributions of naturally occurring raindrops, and coupled this information with data from experiments in which they let water droplets fall onto volcanic ash — mimicking the conditions in which the fossil raindrops formed.

Som *et al.* conclude that the atmospheric density 2.7 billion years ago was probably 50 to 105% of that today. This finding immediately calls into question solutions to the faint young Sun paradox that invoke elevated concentrations of greenhouse gases, unless small increases of greenhouse-gas concentration were able to exert a large warming effect. It is also unlikely that higher concentrations of greenhouse-enhancing nitrogen could have caused the paradox, because concentrations of twice or more the present atmospheric abundance would have been required to provide

sufficient warming in the presence of a modest increase in carbon dioxide¹³. Under such conditions, the atmospheric density would have been greater than that predicted by the authors. It therefore seems that elevated concentrations of highly effective greenhouse gases, such as methane¹⁴, ethane¹⁵ and/or carbonyl sulphide¹⁶, may be required to explain the paradox, possibly in combination with moderately higher concentrations of less-effective greenhouse gases such as carbon dioxide¹⁴. A lower planetary albedo, caused by the reduction or absence of continental ice sheets, could also have contributed to warming.

Although raindrop size distributions associated with typical storms are well known, it is possible — albeit unlikely — that the ancient raindrops responsible for the fossilized imprints were unusually large. Small errors in the inferred size of the raindrops would result in significant errors in the atmospheric pressure predicted by Som and colleagues' method³. The accuracy of the method is further limited by lack of information about

W. ALTERMANN/UNIV. PRETORIA

factors (such as moisture content) that would have affected the cohesiveness of the ash in which the fossil imprints were made; the cohesiveness can affect the morphology of impact craters. The atmospheric density 2.7 billion years ago could therefore have been more than twice that of the modern atmosphere if the circumstances under which the fossil imprints formed were unusual.

It is to be hoped that Som and colleagues' work will stimulate further studies of fossil raindrop imprints, including perhaps those originally observed by Lyell. In particular, it will be interesting to see whether coherent temporal trends in atmospheric pressure can be inferred from imprints in deposits of varying ages. With increasing recognition and analysis of such features in the geological record, it may be possible to establish a

chronological record of atmospheric pressure on Earth throughout the past 3.5 billion years. Such a record would shed light on otherwise poorly constrained aspects of climate change deep in Earth's history. ■

William S. Cassata and Paul R. Renne are in the Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, California 94720, USA, and at the Berkeley Geochronology Center.
e-mails: cassata@berkeley.edu; prenne@bgc.org

1. Ribas, I. in *Solar and Stellar Variability: Impact on Earth and Planets* (eds Kosovichev, A. G., Andrei, A. H. & Rozelot, J.-P.) Ch. 1 (Cambridge Univ. Press, 2010).
2. Sagan, C. & Mullen, G. *Science* **177**, 52–56 (1972).
3. Som, S. M., Catling, D. C., Harnmeijer, J. P., Polivka, P. M. & Buick, R. *Nature* <http://dx.doi.org/nature10890> (2012).
4. Ramsay, J. G. *Trans. Geol. Soc. S. Afr.* **66**, 353–401 (1963).
5. Schopf, J. W. & Barghoorn, E. S. *Science* **156**, 508–512 (1967).
6. Harrison, T. M. *Annu. Rev. Earth Planet. Sci.* **37**, 479–505 (2009).
7. Reinhard, C. T. & Planavsky, N. J. *Nature* **474**, <http://dx.doi.org/10.1038/nature09959> (2011).
8. Dauphas, N. & Kasting, J. F. *Nature* **474**, <http://dx.doi.org/10.1038/nature09960> (2011).
9. Goldblatt, C. & Zahnle, K. J. *Nature* **474**, <http://dx.doi.org/10.1038/nature09961> (2011).
10. Rosing, M. T., Bird, D. K., Sleep, N. H. & Bjerrum, C. J. *Nature* **474**, <http://dx.doi.org/10.1038/nature09962> (2011).
11. Lyell, C. *Edinb. N. Phil. J.* **51**, 70–74 (1851).
12. Kasting, J. F. *Precamb. Res.* **34**, 205–229 (1987).
13. Goldblatt, C. et al. *Nature Geosci.* **2**, 891–896 (2009).
14. Sheldon, N. D. *Precamb. Res.* **147**, 148–155 (2006).
15. Haqq-Misra, J. D., Domagal-Goldman, S. D., Kasting, P. J. & Kasting, J. F. *Astrobiology* **8**, 1127–1137 (2008).
16. Ueno, Y. et al. *Proc. Natl Acad. Sci. USA* **106**, 14784–14789 (2009).